UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

UN MODÈLE DE SIMULATION POUR UN SYSTÈME D'URGENCES SANTÉ : LA MÉTHODOLOGIE

par

Pierre Trudeau
Jean Choquette
Jacques Hodgson
Jacques Ferland
Jean-Marc Rousseau

Cette recherche a été rendue possible grâce à une subvention de F.C.A.R. dans le cadre du programme "Actions spontanées".

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal
Centre de recherche sur les transports - Publication #530
février 1987

Dépôt légal - 3e trimestre 1987

TABLE DES MATIÈRES

Page

RÉSUMÉ.....iii

CHAPITRE 1 : INTRODUCTION.....1

CHAPITRE 2 : MÉCANIQUE GÉNÉRALE DE LA SIMULATION.....4
2.1 - Définition.....4
2.2 - Les avantages et les désavantages de la simulation.....4
2.3 - La simulation dynamique par événements discrets.....5
2.4 - Les notions d'état et d'événement.....7
2.5 - L'horloge.....8
2.6 - La notion de décision.....8
2.7 - Les différentes composantes d'un modèle de simulation.....9
2.8 - Les qualités d'un bon modèle.....9

CHAPITRE 3 : INFORMATIONS NÉCESSAIRES POUR LA DESCRIPTION DU SYSTÈME.....12
3.1 - Généralités.....12
3.2 - Caractéristiques du système subies par le planificateur.....12
3.2.1 - Les données de base.....12
3.2.2 - Les variables aléatoires.....13
3.3 - Données sur lesquelles on peut agir.....15

CHAPITRE 4 : APERCU GÉNÉRAL DU MODÈLE DÉVELOPPÉ.....18
4.1 - Aperçu général.....18
4.1.1 - Le choix du langage de programmation.....18
4.1.2 - Les composantes du système.....19
4.2 - Description des événements.....20
4.3 - Les paramètres de base du modèle.....21

CHAPITRE 5 : LE GÉNÉRATEUR DES DEMANDES.....24
5.1 - Généralités.....24
5.2 - Description des variables générées.....24

CHAPITRE 6 : LE SIMULATEUR.....28
6.1 - Aperçu général.....28
6.2 - Les variables et structures de données globales.....28
6.3 - Les procédures et fonctions.....33
6.3.1 - Initialisations et traitements.....33
6.3.2 - Procédures de gestion de l'horloge.....34
6.3.3 - Procédures relatives à chacun des événements.....35
6.3.4 - Procédures et fonctions pour le déplacement des ambulanciers.....36

TABLE DES MATIÈRES

Page

CHAPITRE 7 : LE MODULE DE DÉCISION.....40
7.1 - Généralités.....40
7.2 - Description des tactiques.....41
7.2.1 - Urgences Santé no 1.....41
7.2.2 - Urgences Santé no 2.....45
7.2.3 - Garage.....47

CHAPITRE 8 : LE MODULE D'ANALYSE DES RÉSULTATS.....49
8.1 - Généralités.....49
8.2 - Contexte horizon fini.....51
8.3 - Contexte horizon infini.....53
8.4 - Autres procédures statistiques.....61

CHAPITRE 9 : LE SYSTÈME D'INTERFACES.....64
9.1 - Généralités.....64
9.2 - Description de la structure de fichiers.....65
9.2.1 - La structure globale.....65
9.2.2 - Description des fichiers de paramètres résidants sur le F.P. de base.....67
9.2.3 - La banque de postes d'attente.....71

CHAPITRE 10 : LE DÉVELOPPEMENT DU MODÈLE.....72
10.1 - Le cas particulier : Urgences Santé.....72
10.2 - La collecte de données.....72
10.2.1 - Les données de base.....73
10.2.2 - Les variables aléatoires.....76
10.3 - La validation.....78
10.3.1 - Définitions.....78
10.3.2 - Considérations sur la validation.....79
10.3.3 - Une approche en trois étapes.....80

CHAPITRE 11 : CONCLUSION.....83

RÉSUMÉ

Nous présentons un modèle de simulation d'une flotte de véhicules ambulanciers (système d'Urgences santé). Il s'agit d'une simulation par événements discrets permettant de reproduire dans le temps les opérations du système, afin d'en évaluer les différentes performances (temps de réponse, charges de travail des ambulances, etc.). Plusieurs tactiques d'opération (avec ou sans relocalisation des véhicules) peuvent être étudiées sur un horizon fini ou infini.

La modélisation du système est expliquée en détail. La génération des variables aléatoires et l'analyse statistique des résultats de simulation sont traitées avec un soin particulier et une méthodologie de validation du modèle est présentée.

Mots-clés : système d'urgence, simulation, évaluation de tactique d'opération, analyse statistique, validation.

ABSTRACT

In this paper, we introduce a simulation model of an ambulance fleet — an emergency system from Urgences santé. It is a discrete event simulation which allows the generation of system operations over a period of time in order to evaluate the different levels of performance (response time, workload of ambulances, etc.). Many operation tactics, with or without relocation of vehicules, are analyzed on a finite or infinite horizon.

We explain in detail the system modelization, and carefully examine the generation of the random variables and the statistical analysis of the results of a simulation. In addition, we describe a validation methodology for such a system.

Key words : emergency system, simulation, operation tactic evaluation, statistical analysis, validation.

CHAPITRE 1
INTRODUCTION

La planification et l'opération d'un système d'Urgences santé suscitent des problèmes considérables. Dans [1], on présente en détail la problématique qu'elles définissent, ainsi que les différentes solutions pouvant être retenues pour chacun des sous-problèmes. Un de ceux-ci est la génération et l'analyse de tactiques opérationnelles. Les auteurs expliquent qu'il existe un large éventail de contextes possibles, allant du plus contraint au moins contraint, et ils présentent différentes méthodes analytiques et heuristiques pour générer la meilleure tactique possible dans chaque cas [1, section 6.9].

Cependant, lorsqu'on parle de meilleures tactiques, il est nécessaire d'avoir un moyen sûr d'évaluer celles-ci. Ce besoin est d'autant plus critique que nous évoluons dans un contexte complexe où la santé du public est mise en jeu et où il ne saurait être question d'expérimentation qui mettrait des vies en danger. Notre propos, ici, est donc de présenter un outil permettant d'évaluer les performances d'un système d'Urgences santé pour différents contextes donnés (avec ou sans relocalisation, etc.).

Pour en arriver là, il est absolument nécessaire de développer un modèle du système que nous voulons étudier. Un tel modèle devra donner une bonne description du système tout en utilisant un ensemble raisonnable de paramètres. En plus de reproduire son comportement, il devra évaluer différentes mesures des performances d'intérêt (temps de réponse, charges de travail, etc.).

Le système que nous avons à modéliser est complexe et comporte plusieurs éléments aléatoires (comme l'arrivée des appels, le temps de service, le nombre et l'identité des véhicules disponibles). Pour traiter le problème de l'évaluation des performances dans une telle situation, il

existe deux approches : les modèles analytiques et les modèles de simulation.

Les modèles analytiques fournissent une expression mathématique du système dans laquelle il suffit de substituer les différents paramètres le décrivant pour obtenir rapidement et à peu de frais des mesures de performances du système. La formulation mathématique est obtenue par une analyse qui, dans le cas d'un système d'urgence, regroupe des éléments de la théorie des probabilités, des files d'attente et des réseaux. Le degré de sophistication mathématique pour obtenir une telle formulation est très élevé et nécessite de poser plusieurs hypothèses qui sont parfois simplificatrices. Enfin, les modèles analytiques ne permettent en aucun cas de certifier si la modélisation réalisée reflète bien la réalité.

De là découle la pertinence d'un modèle de simulation : il permet de vérifier la vraisemblance des résultats issus d'un modèle analytique. De plus, un tel outil permet de modéliser des situations trop complexes pour être approchées analytiquement. Par contre, la simulation ne permet pas de déterminer des solutions optimales. C'est un laboratoire dans lequel on peut tester des tactiques de déploiement sans avoir à les essayer dans la réalité.

Dans le présent ouvrage, nous présentons un modèle de simulation d'un système d'urgences. Bien que ce modèle se veuille le plus général possible, il a été développé en étroite collaboration avec Urgences santé à Montréal. L'exposé débute avec des considérations sur la mécanique générale de la simulation (chapitre 2) afin de permettre au lecteur non-initié de s'y familiariser. Au chapitre 3, on traite des informations nécessaires à la description du système. Le chapitre 4 donne un aperçu général du modèle développé, alors que les chapitres 6 à 9 traitent plus en détails

des différents modules du système de simulation. Enfin au chapitre 10, nous exposons les différentes étapes du développement du modèle dans le cas particulier d'Urgences santé à Montréal.

CHAPITRE 2

MÉCANIQUE GÉNÉRALE DE LA SIMULATION

2.1 Définition de la simulation

La simulation est un procédé qui permet d'imiter un processus réel. Comme pour les modèles analytiques, il faut poser une série d'hypothèses sur son fonctionnement. Ces dernières prennent souvent l'allure de relations mathématiques. On en arrive ainsi à bâtir un modèle mathématique du système. Un modèle de simulation est dit mathématique en raison de sa formulation et non à cause de sa technique de calcul des performances, car celle-ci ne consiste pas à résoudre un système d'équations quelconque (linéaires ou non linéaires).

Pour évaluer les performances du système, on doit plutôt alimenter le modèle d'une série de données aléatoires (des informations sur les appels et les temps de service dans notre cas) qui reflètent celles obtenues dans la réalité et observer les réactions du modèle face à ces données.

2.2 Les avantages et les désavantages de la simulation

Les avantages

- La simulation permet de modéliser des systèmes complexes (comportant un niveau élevé de détails) ne pouvant être étudiés analytiquement.
- Il est possible de s'assurer de la validité d'un modèle ayant les mêmes caractéristiques d'opération qu'un système existant dans la réalité (technique de validation).
- Elle permet de comparer deux solutions potentielles et d'identifier la meilleure.
- C'est un laboratoire où le temps est condensé. Par exemple, on peut simuler un mois d'opération en 10 000 secondes de temps ordinateur.
- Il est possible de chercher des solutions par essais-erreurs.
- La simulation permet de valider les modèles analytiques.

Les désavantages

- Les modèles de simulation coûtent cher à développer et à utiliser.
- Si le modèle développé n'est pas valide, alors il sera peu utile pour étudier le système réel.
- La simulation est une approche basée sur des expériences. Il peut être difficile de déterminer des relations de cause à effet par une simulation, en particulier lorsque plusieurs paramètres doivent être spécifiés. (Les analyses de sensibilité sont plus compliquées.)
- L'analyse statistique des résultats est difficile pour certains cas (horizon infini).
- Enfin, la simulation ne fournit pas de solutions optimales.

2.3 La simulation dynamique par événements discrets

Il existe plusieurs types de simulation. Notre but n'est pas d'en faire la revue. À ce propos, nous référons le lecteur intéressé à [2]. Disons simplement que, parmi les simulations numériques (utilisant l'ordinateur), on distingue celles qui sont continues de celles qui sont discrètes. Nous ne nous intéresserons qu'aux dernières, et plus particulièrement à celles qui sont dynamiques, i.e. qui évoluent dans le temps.*

Le type de simulation qui retient notre attention est communément appelé simulation par événements discrets (Discrete-Event-Simulation) et s'applique à des systèmes dont l'état change dans le temps.** Ce qui caractérise ces simulations, c'est l'état du système modélisé qui ne change qu'en une quantité dénombrable de moments. Ces moments sont ceux où des événements se produisent comme, par exemple, l'arrivée d'un appel, l'affectation ou le retour en disponibilité d'une ambulance. Le modèle est alimenté en entrée, d'une séquence d'événements (des appels) qui eux, génèrent d'autres événements (affectation, arrivée sur les lieux, retour en disponibilité).

* NOTE : On parle de simulations dynamiques par oppositions aux simulations statiques comme les méthodes de Monte-Carlo par exemple.

**NOTE : Ces systèmes peuvent être discrets ou continus (dans ce cas, ils changent d'états en tout temps).

La simulation consiste à laisser aller ce processus de façon à recueillir les informations permettant d'estimer les performances du système (temps de réponse, charges de travail, etc.).

Nous allons nous intéresser à deux contextes de simulation : horizon fini et horizon infini. Chacun d'eux requiert une approche différente dans l'analyse statistique des résultats. Dans une simulation à horizon fini, les observations sur le système simulé sont recueillies dans un intervalle de temps $[0, T_E]$, T_E étant l'instant où l'événement E se produit. Cet événement E est spécifié avant le début de la simulation. Les mesures obtenues dépenderont explicitement de l'état du système au temps 0. On doit donc prendre soin de bien spécifier les conditions initiales.

Dans une simulation à horizon infini, les performances du système sont définies comme des limites lorsque la durée de la simulation tend vers l'infini. On parle alors d'un processus stochastique à l'équilibre et il n'existe pas d'événement E qui mette fin à la simulation. La durée de celle-ci doit, en pratique, être suffisamment longue pour permettre une bonne estimation des quantités d'intérêt.

Dans la réalité, nous avons le plus souvent affaire à des systèmes ayant une existence limitée. De plus, dans un système d'urgence réel, la demande n'est pas stationnaire, c'est-à-dire que le taux d'appel fluctue dans le temps. Dans un tel contexte, une simulation à horizon fini est évidemment plus appropriée. Toutefois, les simulations à horizon infini s'avèrent d'une grande utilité pour mettre à l'épreuve des modèles analytiques qui utilisent l'hypothèse d'équilibre ("steady state"). De telles simulations permettent en outre de comparer des tactiques sur un même pied d'égalité alors que le taux d'appels est constant comme le supposent les modèles analytiques dont elles sont issues (voir [25]). Cependant, le traitement statistique des résultats en contexte horizon infini est passablement plus difficile (voir chapitre 8).

2.4 Les notions d'état et d'événement

Un bon modèle de simulation doit reposer sur une définition adéquate de l'état du système. Le système que nous considérons (système d'urgence) évolue d'état en état selon des probabilités de transition qui ne dépendent que de son état courant. Pour prédire son comportement futur au temps t, il devrait être suffisant de n'utiliser que l'information du système à cet instant. Ceci rejoint la définition de processus de Markov ([6, p. 205]). Afin d'être consistant avec cette dernière, nous définissons l'état du système d'urgence comme étant :

- le statut et la localisation des ambulances (pour l'affectation et la relocalisation des véhicules);
- la localisation des appels courants (pour le traitement des files d'attente);
- le statut des postes d'attente (pour la relocalisation des véhicules).

Ces informations devraient être suffisantes et connues en tout temps.

Dans la modélisation que nous avons adoptée, les changements d'état du système se produisent lorsque surviennent des événements comme, par exemple, l'arrivée d'un appel, l'affectation d'un véhicule, etc. Dans le déroulement de la simulation, ces événements doivent tour à tour être examinés afin de mettre à jour l'état du système ainsi que pour recueillir les observations sur ses performances. De plus, il arrive fréquemment que l'on doive générer un ou plusieurs événements qui découlent de l'événement courant examiné et qui surviendront plus tard dans la simulation. On distingue donc deux types d'événements :

- ceux qui sont générés aléaltoirement par le modèle, comme l'arrivée d'un appel (ou qui sont fournis en entrée au modèle, comme le début et la fin d'une période de disponibilité d'un véhicule);
- ceux qui découlent de l'arrivée d'événements antérieurs, comme l'affectation, l'arrivée sur les lieux et le retour en disponibilité d'un véhicule.

2.5 L'horloge

Pour gérer le déroulement d'une simulation, il est absolument nécessaire d'avoir un mécanisme qui permette la mise à jour de l'état du système. Puisque les changements d'état surviennent uniquement lorsque des événements se produisent, il n'est pas nécessaire d'examiner le système entre deux événements consécutifs. Il suffit de garder une liste triée des prochains événements à survenir.

La fonction du mécanisme de mise à jour, que l'on appelle horloge, est d'identifier le prochain événement à survenir et de déclencher une procédure qui en fera l'examen. Ce dernier consiste à mettre à jour l'état du système et les compteurs statistiques, à déterminer tous les événements générés par l'événement examiné et à insérer ceux-ci dans la liste des événements à survenir. De plus, il devrait être possible de modifier l'information sur certains événements à survenir, comme par exemple, lors d'une annulation de service où il faut retirer de la liste l'événement "arrivée sur les lieux".

Compte tenu du grand nombre de fois où la liste des événements peut être consultée dans une simulation, il importe de choisir une structure de données et des procédures de mise à jour efficaces pour cette liste. En pratique, on utilise souvent une structure de TAS (voir [8], p. 84).

2.6 La notion de décision

Lorsque surviennent certains événements, une décision doit être prise. Par exemple, lors de l'arrivée d'un appel, on doit décider si celui-ci est placé en file d'attente ou servi immédiatement et, dans ce cas, on doit déterminer quel(s) véhicule(s) envoyer. Lors d'une fin de service, on doit décider où envoyer le véhicule devenu disponible. S'il est possible de faire des relocalisations de véhicules, on doit se demander quand et comment faire celles-ci. On ne peut donc laisser aller la simulation sans

qu'intervienne durant celle-ci un processus de décision. Il est nécessaire de spécifier a priori un ensemble de règles pour chaque événement impliquant une décision.

2.7 Les différentes composantes d'un modèle de simulation

Nous avons fait, jusqu'à présent, un survol de la mécanique générale de la simulation. Nous reportons le lecteur à la figure 1, afin qu'il puisse y visualiser les différentes composantes. D'une manière générale, un simulateur est composé d'un programme principal, d'un bloc d'initialisation, d'une procédure de gestion du temps (horloge), de procédures pour examiner les événements et d'un bloc d'analyse statistique.

2.8 Les qualités d'un bon système de simulation

Les principales qualités attendues d'un système sont les suivantes :

- Reproduire le plus fidèlement possible les opérations du système d'urgence.
- Être suffisamment détaillé afin de fournir des résultats représentatifs de la réalité.
- Ne pas être trop détaillé pour limiter le plus possible les informations à amasser.
- Posséder un bon ensemble de familles de tactiques déjà programmées et facilement "paramétrisables". De plus, il devrait être aisé de définir de nouvelles tactiques.
- Être modulaire dans sa conception pour être facilement modifiable.
- Être facile d'utilisation, i.e.
  - avoir une bonne documentation;
  - posséder une bonne structure de fichiers pour définir les paramètres d'une simulation et assurer les liens entre les différents modules de traitement.

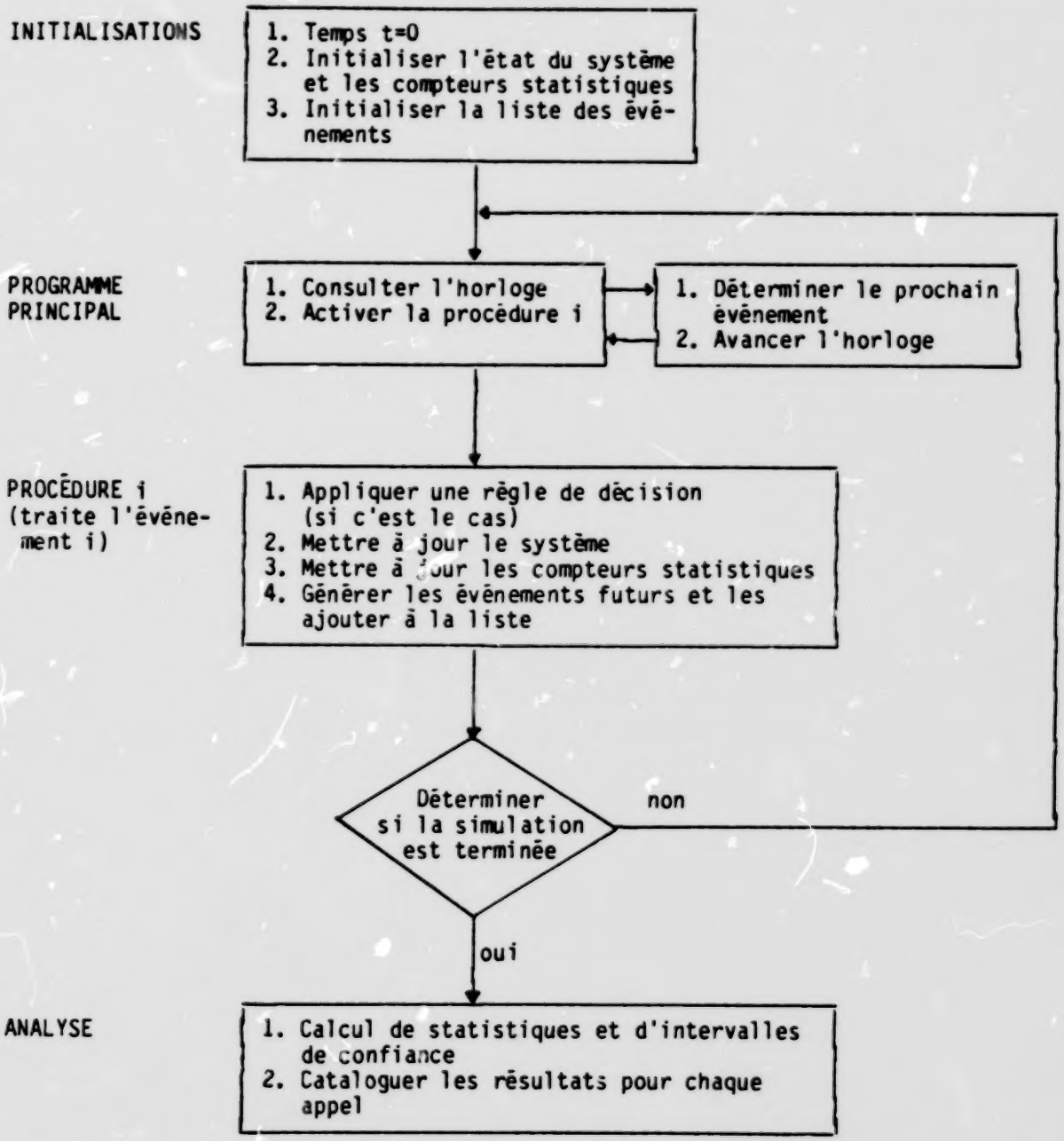

Figure 1 : Schéma d'un simulateur

- Il devrait être possible de développer sans difficulté un système de gestion du modèle (interface usager, base de données, etc.).
- Les programmes d'ordinateur doivent être efficaces, l'espace mémoire requis raisonnable.
- Être facilement transférable, en particulier vers un environnement de micro-ordinateur.
- Être adaptable à n'importe quelle ville et à des systèmes similaires (pompiers, police, etc.).

CHAPITRE 3

INFORMATION NÉCESSAIRE POUR LA DESCRIPTION DU SYSTÈME

3.1 Généralités

Bien que nous ayons voulu développer un modèle de simulation qui puisse s'adapter à différentes situations (ambulance, pompier, police, ...) et à différentes villes, une collecte de données s'avère nécessaire pour chaque cas particulier afin de fixer les valeurs des différents paramètres du modèle. Pour développer ce modèle, nous nous sommes basés sur les données d'Urgences santé à Montréal. Au chapitre 10, nous donnons les détails relatifs à cette analyse particulière. Pour le moment, nous allons nous contenter d'identifier chacun des paramètres et d'en justifier la présence. Nous distinguons deux types d'information : celles qui sont subies par le planificateur et celles sur lesquelles il peut agir (description des règles de décision).

3.2 Caractéristiques du système subies par le planificateur

Dans un système d'urgence, des incidents se produisent un peu partout sur le territoire. Le moment et le lieu exact de chacun d'eux ne sont pas connus à l'avance. Ces incidents donnent lieu à des appels d'urgence. Cette situation est celle qui est subie par le planificateur. Nous définissons maintenant les points qui la caractérisent.

3.2.1 Données de base

Ces données sont connues et fixées pour toute la durée de la simulation. Elles touchent principalement la représentation du territoire et les vitesses. Nous décrivons les éléments nécessaires dans chaque cas.

1) Représentation du territoire

Le territoire est divisé en secteurs afin d'y localiser la demande. Pour chacun des secteurs, on doit déterminer un centroïde (un noeud). C'est de ce dernier qu'émanera la demande durant la simulation.

Attributs du noeud :

- Son numéro
- Table des distances : elle donne la distance du plus court chemin entre chaque paire de noeuds. Cette table permet de calculer les temps de déplacement des véhicules.
- Table des prédécesseurs : pour chaque paire de noeuds a et b, ce tableau donne la liste des successeurs sur le plus court chemin de a à b. Il est nécessaire pour localiser les véhicules en mouvement.

2) Les vitesses

Nous considérons deux vitesses de déplacement : une pour les appels d'urgence (lors de l'affectation et du transport du patient) et une autre pour les déplacements courants (par exemple, lors des retours en disponibilité et des relocalisations). De plus, il nous est possible de prendre en considération les différentes zones de congestion du réseau routier. Dans ce cas, il faut utiliser une structure de réseau pour représenter le territoire dans le but de construire la table des distances entre les secteurs. Des facteurs multiplicatifs de dilatation propres à chacune des zones de congestion sont appliqués à chacun des arcs du réseau. Les distances entre chaque paire de centroïdes de secteurs sont obtenues en calculant tous les plus courts chemins sur le réseau dilaté. Cette technique est basée sur des travaux faits dans le cadre du développement de routes et d'horaires pour le transport adapté des handicapés [L8J]. Une telle procédure nécessite toutefois une analyse de données élaborée afin de codifier le réseau et de calibrer les vitesses sur celui-ci (voir chapitre 10).

3.2.2 Les variables aléatoires

Ces variables désignent tous les éléments non-déterministes du système. Pour chacune d'elles, il faut déterminer sa loi de probabilité et estimer les paramètres qui s'y rapportent. On peut essayer, dans un premier

temps, de trouver une loi de probabilité théorique à laquelle pourrait bien s'ajuster les données. Sinon, dans un deuxième temps, on modélise la fonction de répartition empirique. Dans ce cas, la taille de l'échantillon doit être suffisamment grande (voir [2, chapitre 4] et [3, chapitre 5]. Au chapitre 3, nous verrons comment nous avons modélisé ces lois de probabilité, tandis qu'au chapitre 10, nous discuterons de la collecte des données.

Les variables aléatoires touchent essentiellement deux aspects du système, soit l'arrivée des appels et le service, i.e. la période durant laquelle un véhicule est affecté à un appel.

- Arrivée des appels - les variables suivantes sont nécessaires :
  - le temps où l'appel se produit
  - le numéro du noeud d'où il émane
  - la priorité

    (note : il peut y avoir différents niveaux de priorité - voir chapitre 10 pour le cas de Montréal).

- Service :
  - type de service : c'est une variable binaire qui indique si le service est annulé ou non. Un service annulé indique qu'il n'y a pas de transport à effectuer, mais que le véhicule doit quand même se rendre sur les lieux de l'appel.
  - temps de service : par définition, c'est le temps compris entre l'affectation du véhicule et son retour en disponibilité. Toutefois, les seuls informations à fournir ne concernent que le temps passé sur les lieux et celui passé au C.H. Le simulateur génère lui-même le temps de déplacement (entre l'affectation et l'arrivée sur les lieux, de même qu'entre le départ des lieux et l'arrivée au C.H.).
  - C.H. : on désigne par là, le centre hospitalier où le véhicule doit transporter le patient dans le cas d'un service non annulé.

3.3 Données sur lesquelles on peut agir

Nous désignons ici l'ensemble des décisions pouvant être prises par un planificateur pour faire face à la situation décrite à la section précédente. Les éléments suivants devraient être spécifiés :

- Nombre de ressources en service en tout temps :

  Dans le cas d'une simulation à horizon infini, il suffit de fournir un seul nombre de véhicules; ceux-ci seront en service pour toute la durée de la simulation. Par contre, dans un contexte horizon fini, il faut déterminer pour chaque portion de temps (chaque heure par exemple), le nombre de véhicules en service. Ce qui justifie la présence de l'élément suivant.

- Horaire des véhicules :

  Pour chaque véhicule, il faut spécifier les débuts et les fins de quarts, ainsi que les débuts et les fins de pauses (horizon fini seulement).

  De plus, pour les débuts de quart et les fins de pause, il est nécessaire de donner le noeud où le véhicule devient disponible.

- Postes d'attente :

  Une ambulance en disponibilité est localisée sur un poste d'attente. Il n'est pas nécessaire qu'elle soit toujours associée au même poste. Ces postes peuvent être situés uniquement sur des centroïdes de secteur. On doit en spécifier le nombre, et pour chacun d'eux, les attributs suivants :

  - le numéro de poste
  - le numéro du noeud (centroïde de secteur)
  - la priorité (certains postes peuvent avoir une plus grande importance sur le plan stratégique).

- Tactique de déploiement :

C'est l'ensemble des règles qui permettent de prendre des décisions durant le déroulement de la simulation. Ces règles doivent être programmées. Elles constituent le module de décision du modèle de simulation (voir section 4.6).

La tactique de déploiement touche l'affectation des véhicules et la relocalisation.

1) L'affectation des véhicules

Pour chaque appel, il faut déterminer

- s'il est servi immédiatement ou plus tard;
- quel(s) véhicule(s) est(sont) affecté(s) pour le servir.

De plus, il est nécessaire de mentionner si, une fois les ressources affectées à l'appel, il est possible d'annuler une de celles-ci pour la remplacer par une autre plus près ou pour l'envoyer à un appel plus prioritaire.

2) La relocalisation

La relocalisation se définit comme le changement des emplacements des véhicules en disponibilité. Le but de toute relocalisation est de se donner la meilleure protection possible contre les événements futurs en modifiant adéquatement la manière dont les unités disponibles sont localisées sur le territoire.

Une relocalisation peut survenir dans plusieurs circonstances. D'une manière générale, il peut être avantageux de faire une relocalisation lorsqu'un véhicule disponible devient occupé (arrivée d'un appel, fin de quart, début de pause) et, inversement, lorsqu'un véhicule occupé devient disponible (fin de service, début de quart, fin de pause).

Le problème dans un contexte où il est possible de relocaliser des unités se ramène à identifier les moments opportuns où l'on peut relocaliser et la manière dont cela doit être fait. Dans ce dernier cas, on doit identifier quelles unités sont relocalisées, où et comment elles le sont (c'est-à-dire quelles routes empruntent-elles pour atteindre leurs nouveaux emplacements).

Dans le présent chapitre, nous avons énuméré les informations nécessaires à la construction du modèle. Au chapitre suivant, nous donnons un aperçu général du modèle développé à partir de ces informations.

CHAPITRE 4
APERÇU GÉNÉRAL DU MODÈLE DÉVELOPPÉ

4.1 Aperçu général

4.1.1 Le choix du langage de programmation

Un aspect important dans le développement d'un modèle de simulation est le choix de l'environnement informatique et, plus particulièrement, du langage de programmation. Il existe des langages propres à la simulation qui sont très répandus, tels que SIMSCRIPT et GPSS. L'avantage de ces langages est qu'on y trouve un ensemble de procédures pré-définies qui permettent d'écrire des programmes de simulation d'une manière plus concise et plus transparente. Toutefois, pour développer un outil de simulation comportant un niveau élevé de détails, ces langages ne sont pas recommandés [2, p. 215]. La principale raison est qu'ils apparaissent comme des boîtes noires dont le comportement interne est difficilement compréhensible par l'usager. Il est alors préférable d'utiliser un langage général de programmation dont le fonctionnement nous est familier de façon à développer un système dont il sera possible de prédire le comportement dans ses moindres détails. C'est ce que nous avons fait.

Bien sûr, il existe des langages de simulation qui sont également de puissants langages de programmation (comme SIMULA). Toutefois, leur usage est moins répandus. Nous avons plutôt opté pour le langage PASCAL qui permet une bonne structuration des données, une approche modulaire et l'allocation dynamique de la mémoire, ce qui est important pour la programmation des files d'attente. De plus, PASCAL étant assez répandu, le modèle pourra être implanté sur d'autres ordinateurs. Nous avons utilisé la version PASCAL 6000 sur un ordinateur CDC CYBER [16].

4.1.2 Les composantes du système

Le déroulement d'une simulation s'effectue en plusieurs étapes. Le modèle est constitué d'un ensemble de composantes permettant d'exécuter chacune d'elles. Ces étapes sont les suivantes :

1) Fixer les paramètres globaux de la simulation (voir section 4.3).

2) Créer un fichier d'appels.

3) Sélectionner la tactique d'opération de même que les paramètres qui lui sont propres.

4) Faire exécuter le programme de simulation proprement dit. Un fichier de résultats est créé (il contient des informations pour chaque appel simulé).

5) Faire exécuter différents programmes d'analyses statistiques sur le fichier de résultats.

La création d'un fichier d'appels est particulièrement utile lorsque l'on compare plusieurs tactiques d'opération. Cela permet d'éviter de générer plusieurs fois les mêmes variables et d'éliminer du même coup les problèmes de synchronisation lorsque deux tactiques sont comparées en couplant les appels pour réduire la variance (voir [2, p. 32-33] et [3, p. 352]). La génération d'un fichier de résultats permet de faire toutes les analyses statistiques désirées même celles auxquelles on n'aurait pas pensé avant l'exécution de la simulation.

Nous présentons brièvement les composantes de l'outil de simulation qui seront décrites en détails dans les prochains chapitres.

1) Le générateur des demandes : crée le fichier d'appels.

2) Le simulateur : effectue la simulation à partir des paramètres lus en entrée et du module de décision décrivant la tactique utilisée.

3) Le module de décision : pour chaque famille de tactiques, ce module contient les procédures qui sont activées par des événements impliquant une décision.

4) Le module statistique : permet de faire l'analyse statistique à partir du fichier de résultats (statistiques descriptives, graphiques, intervalles de confiance, test). Il contient également des procédures pour déterminer la longueur des simulations en fonction d'un degré de fiabilité statistique désiré.

5) Les interfaces : c'est le système de fichiers qui permet de définir les paramètres de la simulation et de faire le lien entre les différentes composantes du modèle.

La figure 2 représente les composantes :

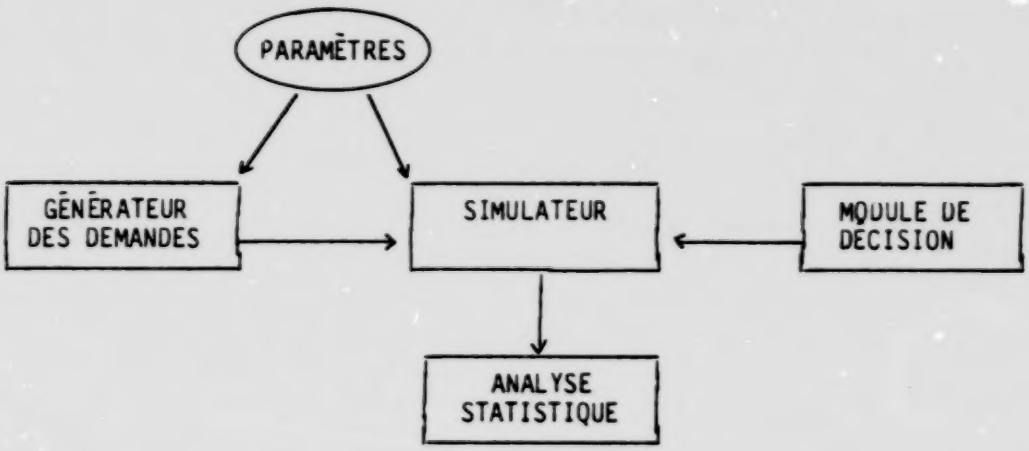

Figure 2 : Les différents modules du modèle.

4.2 Description des événements du modèle

Le tableau 1 donne la description des différents événements propres à notre modèle. Pour chacun d'eux, ce tableau indique également s'il est généré par un événement antérieur, s'il génère des événements futurs et s'il implique une décision.

Tableau 1 : Les différents évènements d'une simulation.

TYPE D'ÉVÈNEMENT	DESCRIPTION	ÉVÈNEMENTS ANTÉRIEURS	ÉVÈNEMENTS GÉNÉRÉS	DÉCISION
TNC	arrivée d'un appel	—	TSS	non
TSS	affectation d'un véhicule	TNC	TSL,TRE	oui
TSL	arrivée du véhicule sur les lieux de l'appel	TSS	TSF	non
TSF	fin de service (retour en disponibilité du véhicule	TSL	TRE	non
TRE	relocalisation des véhicules disponibles	TSS,TSL,TDQ, TFQ,TDP,TFP	—	oui
TDQ	début d'un quart de travail pour un véhicule	—	TRE	non
TFQ	fin d'un quart de travail	—	TRE	non
TDP	début d'une pause pour un véhicule	—	TRE	non
TFP	fin d'une pause	—	TRE	non

Note : Les évènements TNC, TDQ, TFQ, TDP et TFP sont générés en entrée.

4.3 Les paramètres de base du modèle

Dans une expérience de simulation, on doit toujours spécifier un ensemble de paramètres qui déterminent exactement la situation à simuler. Nous ne traitons, ici, que des paramètres globaux, i.e. qui sont communs à toutes les tactiques. Ceux étant propres à chacune de ces dernières seront décrits au chapitre 7.

1) Le type de simulation: horizon fini ou horizon infini.

2) Si le type est horizon fini alors les éléments suivants sont requis :

   - le temps du début de la présimulation et/ou les conditions initiales (voir section 8.2 sur l'analyse d'une simulation à horizon fini);
   - le temps du début de la période simulée;

- le temps à la fin de la période simulée;
- le taux d'appels pour chaque heure de la période simulée;
- le nombre maximal de répétitions (voir chapitre 8).

Sinon (horizon infini) on doit spécifier

- la durée maximale de la simulation;
- le taux d'appels (qui est constant pour toute la durée de la simulation).

3) Le type de temps de service

Essentiellement, le temps de service est de deux types:

1) Global: dans ce cas, un seul temps de service est généré. Il représente le délai entre l'arrivée sur les lieux et le retour en disponibilité. On doit de plus spécifier le type de loi de probabilité (empirique ou exponentielle).

2) Contexte réel: ici, le temps de service est la somme du :
   - temps passé sur les lieux de l'appel;
   - temps de déplacement entre les lieux et le CH;
   - temps passé au CH.

   Les temps passés sur les lieux et au CH sont générés à partir des fonctions de répartition empiriques obtenues des données.

Note: on utilise le type global pour simuler des temps de service dont les lois de probabilité se rapprochent de celles utilisées dans les modèles analytiques. Cela permet d'une part de vérifier la mécanique de la simulation pour des tactiques modélisées analytiquement, et d'autre part, d'étudier la robustesse de ces modèles analytiques en s'éloignant graduellement des hypothèses qu'ils utilisent.

4) Les germes (seeds)

Afin d'initialiser le générateur de nombres aléatoires (chapitre 5) il est nécessaire de spécifier deux germes qui sont en fait des entiers positifs.

5) Les vitesses

Deux fonctions de vitesse sont nécessaires pour le modèle, soient celle pour les déplacements urgents et celle pour les déplacements courants (voir chapitre 10).

6) L'horaire des véhicules

Dans une simulation à horizon infini, on spécifie le nombre de véhicules. Ces derniers seront en disponibilité pour toute la durée de l'expérience. Dans un contexte horizon fini, il faut identifier, pour chaque véhicule, le début et la fin des quarts de travail, de même que le début et la fin des pauses prévues (s'il y a lieu).

Dans les chapitres suivants (6 à 9), nous abordons en détails chacune des composantes du système, soient le générateur des demandes, le simulateur, le module de décision, le module d'analyse statistique et le système d'interfaces.

CHAPITRE 5

LE GÉNÉRATEUR DES DEMANDES

5.1 Généralités

Le but du générateur des demandes est de créer toutes les informations relatives à chacun des appels (arrivée et service). Ces informations sont cataloguées sur un fichier qui est utilisé en entrée avec le simulateur. Le générateur utilise les paramètres globaux décrits au chapitre précédent.

Note sur le générateur de nombres aléatoires

Afin de générer des observations provenant d'une loi de probabilité donnée, nous utilisons toujours des nombres pseudo-aléatoires issus d'une loi uniforme [0,1). Pour générer ces nombres, nous utilisons le bloc RANDOM disponible sur PASCAL 6000 [16].

5.2 Description des variables générées et des méthodes utilisées

Le temps de l'appel

Nous supposons que l'arrivée des appels se comporte comme un processus de Poisson (au chapitre 10, nous discuterons de la validité de cette hypothèse), chaque noeud étant une source indépendante d'appels. Dans un contexte horizon infini, on considère un processus de Poisson(λ) constant où λ est le taux d'appel (nombre moyen d'appels par minute). La technique consiste à générer une séquence de temps inter-arrivées issus d'une loi exponentielle dont l'espérance est $1/\lambda$ [3, p.270].

Note : on génère une variable x issue d'une loi exponentielle en utilisant la transformation inverse [3, p.254]:

1) générer $u \sim U(0,1)$
2) $x \leftarrow -1/\lambda \ln u$.

Dans un contexte horixon fini, on considère plutôt un processus de Poisson non-constant. Dans un tel processus, le taux d'appel dépend explicitement du temps. Il est modélisé comme une fonction $\lambda(t)$ définie sur la période $[T_0, T_E]$, où t désigne le temps. Il est alors possible de générer une séquence de moments d'arrivée. Nous avons employé la méthode de la transformation inverse [17]. Les étapes sont les suivantes:

1) $i \leftarrow 0$; $T_i \leftarrow T_0$; $S_i \leftarrow 0$

2) Tant que $T_i < T_E$ faire

$i \leftarrow i+1$

générer $u_i \sim U(0,1)$

$S_i \leftarrow S_i - \ell n(Ui)$

$T_i \leftarrow \Lambda^{-1}(S_i)$

Note: $\Lambda(T) = \int_{T_0}^{T} \lambda(t)dt$

Le lecteur trouvera toutes les justifications théoriques dans la référence [17, p.98-99]. Pour définir la fonction $\lambda(t)$, on spécifie en pratique la période $[T_0, T_E]$, de même que le nombre moyens d'appels pour chaque heure de cette période.

Le noeud d'où émane l'appel

Les appels ne sont générés qu'à partir des centroïdes de secteur (noeuds). Soient n noeuds et soit p(1), p(2),...,p(n), la fonction de masse des appels. Afin de générer le noeud de l'appel, on utilise la transformation inverse décrite dans [3], p.264:

1) Générer $u \sim U(0,1)$

2) Trouver l'entier I satisfaisant

$$\sum_{j=1}^{I} p(j) < u < \sum_{j=0}^{I+1} p(j)$$

I devient le numéro du noeud à générer.

Note: i) afin d'optimiser la recherche en (2), on utilise la séquence des p(j) triée dans l'ordre décroissant;

ii) la fonction de masse p est la même dans les deux contextes (horizons fini et infini);

iii) cette méthode est utilisée pour toutes les variables aléatoires discrètes.

La priorité de l'appel

Il suffit de connaître le nombre de priorités de même que les probabilités associées à chacune d'elles et d'appliquer la méthode décrite ci-haut.

Le type de service

Il est généré à partir d'une loi de Bernoulli (annulé et non-annulé).

Le CH de la destination

Une fois le véhicule arrivé sur les lieux de l'appel, on considère sa destination. S'il y a un transport à effectuer, alors cette destination est un centre hospitalier(CH). Dans le cas contraire (annulation de service), l'ambulance redevient disponible sur place. Dans le cas où le CH doit être généré, on peut procéder de trois façons:

1) prendre le CH qui est le plus près du noeud de l'appel
2) générer le CH à partir de la fonction de masse globale des CH (c'est-à-dire pour tout le territoire)
3) générer le CH à partir de la fonction de masse du CH conditionnelle au lieu de l'appel. Par exemple, pour un appel dans une région A, on considère

$$P[\text{CH.} \mid \text{l'appel est dans la zone A}].$$

Le territoire est préalablement divisé en régions. Leur nombre ne doit pas être trop grand afin de permettre une bonne estimation des fonctions de masse des CH pour chacune d'elles (voir section 10.1 pour le cas de Montréal).

Le temps de service

Si le type de temps de service est global, alors un seul temps est généré. On peut utiliser une fonction de répartition exponentielle [3, p.254] ou la fonction de répartition empirique issue des données [3, p.261]. Dans un contexte réel, on génère un temps sur les lieux et un temps au CH à partir des fonctions de répartition empiriques.

CHAPITRE 6
LE SIMULATEUR

6.1 Aspect général

La principale fonction du simulateur est de faire la gestion de l'horloge. Sa tâche consiste donc à prendre le prochain événement à examiner et à activer la procédure particulière qui en fera l'examen.

Dans ce qui suit, nous allons décrire sommairement le contenu du simulateur, i.e. :

- les variables et les structures de données globales;
- les procédures de gestion de l'horloge;
- les procédures relatives aux événements n'impliquant pas immédiatement une décision (TNC, TSL, TSE, TDQ, TFQ, TDP, TFP).

6.2 Les variables et structures de données globales

Le réseau

- Table des distances entre chaque paire de noeuds (COURTCHEMIN)
  Note: Sous l'hypothèse que le problème est symétrique, (d(i,j)=d(j,i)) la moitié de la table est gardée en mémoire (d(i,j)i<j). La diagonale peut être non nulle (distance intra-secteur).
- Tableau donnant les prédécesseurs sur le p.c.c. entre chaque paire de noeuds (NOEUDSUIVANT).
- Noeuds : donne les coordonnées (x,y) pour chaque noeud du réseau afin de permettre d'identifier le répartiteur associé à un appel (voir la description des files d'attente un peu plus loin).

Les postes d'attente

NBPOSTE : Nombre de postes d'attente (paramètre)

TPOSIT : Noeud de chacun des postes (paramètre)

AMBPOSTE : Pour chaque poste, donne la liste des ambulances disponibles qui y sont affectées (variable).

OCCUPE : Booleen = vrai si le poste est occupé (ou en voie d'être occupé) par une ambulance (variable).

Les files d'attente

Il est possible de tenir compte de la présence de deux répartiteurs. Le but est d'associer un temps non-nul au travail du répartiteur et ainsi de rendre la simulation encore plus réaliste. Ce délai fait partie du temps de réponse et il est bon d'en tenir compte explicitement.

Le territoire est alors divisé en deux régions. Chaque répartiteur s'occupe d'affecter et de relocaliser les ambulances pour la région dont il est responsable. Une structure de file d'attente est associée à chacun des répartiteurs et à chacune des priorités (voir figure 3). Cette structure est décrite par les éléments suivants :

```
TYPE
IDENTREPAR = (EST, OUEST) Type scalaire permettant d'identifier les
                          répartiteurs
TABREPTR   = Tableau [IDENTREPAR] de TABPTR
TABPTR     = Tableau [0..MAXPRIO-1] de PTRLENR
             où MAXPRIO = nombre maximal de priorité
PTRLENR    = ↑LENREGIS
LENREGIS   = Struct
             SUCC : PTRLENR; (pointe sur le prochain élément de la
                              liste)
             ENR  : ENREGIS; (enregistrement contenant l'appel à
                              placer en attente)
             Fin;
VAR
PREMIER, DERNIER : TABREPTR;
```

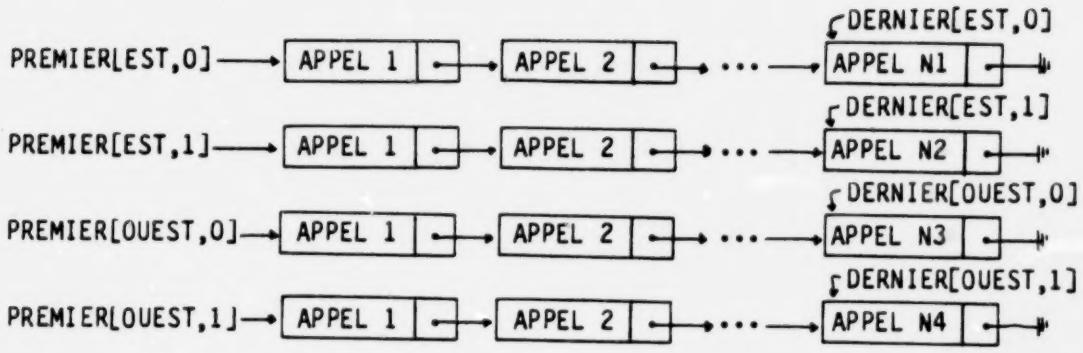

Figure 3 : Structure des files d'attente
(deux priorités)

Les ambulances

NBAMBUL : Nombre maximal d'ambulances considérées dans une même simulation (parmètre);
Horizon infini : nombre d'ambulances en service
Horizon fini : nombre total d'ambulances pour l'ensemble de la simulation.

NAMBDISP : Nombre d'ambulances en disponibilité, i.e. en service et libre pour répondre à un appel (variable).

DISP : Tableau indiquant l'état de chacune des ambulances (variable)
Codes possibles :

0 : disponible

1 : se rend vers les lieux d'un appel

2 : l'ambulance est sur les lieux de l'appel ou en route vers le C.H. ou au C.H.

3 : en dehors de ses heures de service

AMBULANCE : Tableau contenant toutes les informations relatives aux déplacements de chaque ambulance (variable).

- POSDEPART : position* de départ
- POSARRIVE : position* d'arrivée
- HRDEPART : temps lors du départ
- HRARRIVEE : temps lors de l'arrivée
- LCHEMIN : contient la liste des successeurs du dernier chemin emprunté par l'ambulance.
- POSTASSOCIE : poste sur lequel l'ambulance est localisée ou en voie de l'être.

OCCUPATION : Tableau dont chaque entrée pointe sur les informations décrivant l'appel servi par une ambulance lorsque celle-ci est affectée à un appel (variable).

FQIMMINENTE : Tableau indiquant si la fin de quart de chaque ambulance est imminente, i.e. si la fin de quart va survenir dans les quinze prochaines minutes. Cela permet d'entreprendre une action en tenant compte qu'une ressource deviendra bientôt non disponible.

QUARTFINI : Tableau indiquant si le quart de chaque ambulance est terminé (variable).

PAUSE : Tableau indiquant s'il y a une pause associée à chacune des ambulances (variable).

---

* Note : une position sur le réseau est définie par deux noeuds n_1 et n_2, ainsi que par la distance à parcourir pour atteindre le noeud n_2.

L'horloge (TAS) : On utilise une structure de TAS. Celle-ci est représentée dans une zone de mémoire contigüe par le tableau RESULT.

RESULT : Tableau [1...MAXTAS] de ENREGIS.

ENREGIS = struct

- contient toutes les informations relatives à un évènement, quel qu'il soit.

Les informations suivantes sont toujours nécessaires:

TMS : temps de l'évènement

EVENMT : type de l'événement (voir tableau 1, p. 21)

NAMB : numéro de l'ambulance impliquée dans l'évenement

REP : répartiteur associé à l'appel.

Les champs supplémentaires sont utilisés pour conserver les informations sur les appels :

Informations de base fournies en entrée

TAPPEL : temps de l'appel

PRIO : priorité de l'appel

POSIT : noeud d'où émane l'appel

ARR : code de destination (C.H.)

DEST : noeud de la destination

TSERV : temps de service (lorsque global)

DELAILIEU : temps passé sur les lieux (contexte réel)

DELAICH : temps passé au C.H. (contexte réel)

Informations cumulées durant le déroulement de la simulation

TAFFECT : temps à l'affectation

TSURLIEU : temps à l'arrivée sur les lieux

TFINAL : temps lors du retour en disponibilité.

Pour chaque événement, les informations de base sur les appels sont toujours connues.

Autres variables globales permettant la mise à jour du système:

TEMPS : temps courant de la simulation (nombre réel de minutes)

CODE : type de l'événement courant (celui qui est traité)

ENREG : variable de type ENREGIS qui donne toutes les informations sur l'événement courant à traiter. Si l'événement courant génère un autre événement, alors ENREG est remis dans le TAS avec TMS = temps où l'événement sera généré et
EVNMT = nom de l'événement à générer
NAMB = numéro de l'ambulance impliquée dans l'événement à générer.

LONHP : indique combien il y a d'événements dans le TAS.

6.3 Les procédures et fonctions

Nous allons décrire ici les principales procédures du simulateur. Nous ne décrivons cependant pas en détail la liste des paramètres de chaque procédure.

6.3.1 INITIALISATIONS et TRAITEMENT

INITIALISATIONS : fait toutes les initialisations concernant les variables de la simulation qui ne dépendent pas de la tactique.

TRAITEMENT : c'est la procédure de base du système. Son en-tête ne comprend aucun paramètre. TRAITEMENT procède en deux étapes :

1) Elle appelle la procédure MINIMUM pour prendre le prochain événement à traiter et mettre à jour le temps courant de la simulation.
2) Elle active la procédure pour traiter l'événement courant. La structure est la suivante :

```
CAS       CODE DE
TNC             : NOUVAPPEL (        );
TSS             : DEBSERV (          );
TSL             : SURLLIEUX (        );
TSF, TDQ        : DEVIENSLIBRE (     );
TFQ             : FINQUART (         );
TDP             : DEBUTPAUSE (       );
TFP             : FINPAUSE (         );
TRE             : RELOCALISER (      );
TL              : DEBUT
                    MEMORISER (      );
                    FINSIMUL (       );
                  : FIN;
FIN;
```

6.3.2 Procédures pour la gestion du TAS

MINIMUM : retourne dans ENREG les informations concernant le prochain événement à examiner et met à jour le TAS.

Pour des événements se produisant au même moment, on prend ceux-ci dans l'ordre suivant :

TNC, TSS, TSL, TSF, TFP, TFQ, TDQ, TDP, TRE et TL.

TRIER : insère un événement et met à jour le TAS.

TROUVER : permet de trouver et d'enlever un événement du TAS tout en mettant à jour le TAS.

6.3.3 Procédure pour chaque événement

NOUVAPPEL : traite l'arrivée d'un appel (TNC). Le répartiteur associé au noeud de l'appel est identifié. L'appel est placé en file d'attente et un TSS est généré pour lui affecter un véhicule.

Note : Si le répartiteur est libre, le TSS aura lieu au même moment que le TNC.

DEBSERV : affecte une ambulance selon la politique du plus près. Génère un événement TSL (arrivée sur les lieux).

SURLLIEUX : cette procédure est appelée lors de l'arrivée d'un véhicle sur les lieux (TSL). Un délai de service est alors calculé. Il dépend du paramètre TYPTSERV (voir définition des paramètres). À la fin du service, le véhicule redevient disponible au C.H. (sauf dans le cas de la tactique garage où il peut redevenir disponible au garage). L'événement TSF (fin de service) est généré.

DEVIENSLIBRE : traite le cas d'un véhicule devenant disponible à la suite d'une fin de service (TSF) ou d'un début de quart (TDQ). L'état du système est en partie mis à jour. Un événement TRE (relocalisation) est généré afin de relocaliser le véhicule devenu disponible (et peut être d'autres déjà disponibles).

FINQUART : appelée lors d'une fin de quart (TFQ). Des statistiques sur l'ambulance terminant son quart sont compilées. L'état du système est mis à jour et un événement

TRE est généré pour relocaliser les véhicules disponibles, si nécessaire.

DEBUTPAUSE : met un véhicule hors de service. Avant la simulation, on spécifie l'horaire des pauses. NOAMB est l'ambulance devant être affectée à une pause.
Pour la validation du modèle :
Si NOAMB n'est pas libre, alors la pause sera attribuée à l'ambulance disponible la plus près de NOAMB (s'il en existe). Les événements TRE et TFP sont générés.
En production :
La pause est remise jusqu'au retour en disponibilité de NOAMB.

FINPAUSE : rend disponible l'ambulance NOAMB si celle-ci était affectée à une pause. L'état du système est mis à jour et un TRE est généré.

RELOCALISER : cette procédure fait partie du module de décision. Elle est appelée lors d'un TRE. Elle dépend de la tactique utilisée (voir module de décision chapitre 7).

FINSIMUL : est appelée à la fin de la simulation. Elle permet de traiter les événements qui restent dans le TAS sans en générer d'autres. Ces événements ne sont pas comptabilisés dans les statistiques.

6.3.4 Fonctions et procédures relatives aux déplacements des ambulances

FONCTIONS

DISTANCE : calcule la distance entre deux positions

CONVTEMPS : convertit une distance en temps (en minutes). On utilise deux vitesses :

VITESSE 0 : en urgence (à l'affectation et au transport du patient)

VITESSE 12 : autrement.

CONVDIST : convertit un temps de déplacement en distance (en mètres).

PROCEDURES

FAIRECHEMIN : mémorise le chemin emprunté par une ambulance lors d'un de ses déplacements. Le chemin est défini par la liste des successeurs entre la position de départ et celle d'arrivée. Le chemin est, en pratique, représenté par une liste dynamique dont chacune des cellules contient un noeud, ainsi que le temps auquel l'ambulance sera localisée sur ce noeud.

CALCULPOSITION : calcule les positions d'une ambulance, que celle-ci soit immobile ou en mouvement. On utilise la liste dynamique décrite pour la procédure FAIRECHEMIN.

POSITIONNE : Cette procédure est appelée lorsqu'une ambulance disponible doit être déplacée afin de couvrir un poste vacant. Elle effectue le changement de position des ambulances disponibles, i.e. :

- identifie les ambulances qui seront impliquées dans le repositionnement
- détermine le chemin que chacune d'elles empruntera
- met à jour AMBULANCE (pour le déplacement des ambulances) et OCCUPE (pour le statut des postes d'attente).

Un repositionnement doit se faire en un minimum de temps et en impliquant un nombre minimal d'ambulances. On procède en faisant des collisions.

Plus précisément, au début du processus, on construit un petit réseau dont l'ensemble des noeuds est constitué :

- des postes d'attente occupés de priorité* égale ou supérieure à celle du poste à couvrir (où à celle du poste de l'ambulance à déplacer lorsque celle-ci est sur un poste);
- du noeud le plus près de la position de l'ambulance à déplacer. Ce noeud est appelé ici noeud 0.

L'ensemble des arcs est, pour sa part, déterminé de la façon suivante :

- tous les arcs liant chaque paire de noeuds sont triés dans l'ordre croissant de leur longueur;
- des arcs sont ajoutés jusqu'à ce que le noeud à couvrir et le noeud 0 soient connectés.

On détermine ensuite, sur ce petit réseau, le chemin entre le noeud 0 et le noeud à couvrir qui comporte le plus petit nombre d'arcs. Un algorithme de plus court chemin est utilisé en considérant toutes les longueurs d'arcs égales à 1. Les ambulances couvrant les noeuds de ce chemin sont alors déplacées par "collision".

---

*NOTE : Dans une tactique, il est possible de spécifier des postes d'attente de différentes priorités. Les postes prioritaires doivent être couverts en premier. POSITIONNE peut tenir compte de deux niveaux de priorité.

POSITIONNE est utilisée avec toutes les tactiques permettant des relocalisations. Elle peut être appelée à tout moment de la simulation. Il suffit de spécifier le numéro d'une ambulance à relocaliser et le poste qui doit être couvert.

La programmation de chacune des tactiques ne fait pas partie du simulateur. Nous les avons séparés de façon à ce qu'elle constitue un module à part. Au chapitre suivant, nous décrivons ce module.

CHAPITRE 7
MODULE DE DÉCISION

7.1 Généralités

Le module de décision est constitué d'un ensemble de blocs de procédures et fonctions. Chacun de ces blocs correspond à une famille de tactiques paramétriques. La sélection d'une tactique doit donc se faire en choisissant le bloc désiré et en fixant, pour celui-ci, les paramètres qui lui sont propres.

Des décisions doivent être prises lorsque survient un événement TSS (affectation de véhicule) ou TRE (relocalisation). Les procédures activées par le simulateur sont alors les suivantes :

TSS : DEBSERV

TRE : RELOCALISER

La procédure DEBSERV fait partie du simulateur. Elle fait appel à la fonction ONLESERT, qui elle appartient au module de décision. La procédure RELOCALISER fait également partie du module de décision. ONLESERT et RELOCALISER sont définies dans chacun des blocs du module.

Nous expliquons maintenant, pour les deux événements TSS et TRE, à quels niveaux se situent les décisions programmées.

TSS - Affectation (ONLESERT)

Lors de l'arrivée d'un appel, les trois points suivants sont pris en considération :

1) Combien d'ambulances faut-il affecter?

   Dans les données que nous avions à notre disposition pour bâtir le modèle, chaque appel ne se rapportait qu'à un seul patient. En conséquence, une seule ambulance est affectée pour chaque appel. De plus, nous ne considérons pas, pour le moment, d'autre type de véhicule que les ambulances.

2) Est-ce que l'appel est servi immédiatement ou plus tard?
La réponse à cette question dépend de la priorité de l'appel et de la disponibilité des ambulances. Ainsi un appel prioritaire est servi immédiatement si des ambulances sont disponibles pour y répondre. Dans le cas d'un appel non-prioritaire, cela dépend de la disponibilité des ambulances et de la règle de décision propre à la tactique. C'est ici qu'intervient la fonction ONLESERT.

3) Si l'appel est servi immédiatement, quelle ambulance est affectée?
Présentement, nous ne considérons qu'une seule politique de sélection : on choisit celle qui est la plus près. Toutefois, d'autres procédures de sélection pourrait être envisagées ("fixed preference", par exemple).

TRE - RELOCALISATION (RELOCALISER)

Un événement relocalisation (TRE) est généré par un événement TSS, TSF, TFQ, TDP ou TFP. Les décisions à prendre sont alors les suivantes :

- Est-il opportun de relocaliser?
- Si oui, alors quelles unités sont relocalisées et de quelle façon?

Les réponses à ces questions sont traitées dans la procédure RELOCALISER.

Dans ce qui suit, nous allons décrire les blocs de tactiques du module de décision. Pour chacun d'eux, nous exposons le principe de base et nous décrivons les paramètres et les principales procédures et fonctions.

7.2 Description des tactiques

7.2.1 Urgence santé no 1

Cette famille de tactiques est inspirée de la manière dont Urgences Santé gérait les opérations de sa flotte d'ambulances avant décembre 1985.

PRINCIPE

On considère deux types de postes d'attente selon leur priorité :

POSTES DE PRIORITÉ ROUGE : ce sont des postes prioritaires, i.e. que ce sont eux qui doivent être couverts en tout premier lieu. En principe, on voudrait qu'ils soient couverts en tout temps. Partant de ce principe, leur nombre et leurs localisations peuvent donc être choisis de façon à satisfaire au minimum les critères de qualité de service dans un contexte déterministe (voir [1]).

POSTES DE PRIORITÉ NOIRE : ce sont des postes de moindre priorité, i.e. qu'il n'est pas nécessaire qu'ils soient couverts en tout premier lieu. Ils sont utilisés afin d'y localiser les ambulances en surplus lorsque tous les postes de priorité rouge sont occupés. Leur nombre et leurs localisations sont déterminés de façon à :

1) augmenter la qualité du service dont le seuil minimum est, en principe, fixé par le nombre et la localisation des postes prioritaires;
2) faciliter les relocalisations lorsque des postes prioritaires se libèrent.

Le but de la tactique est donc de garder toujours couverts les postes de priorité rouge. Lorsqu'il y a des ambulances en surplus, elles sont placées sur des postes de priorité noire afin d'augmenter la qualité du service. Des relocalisations peuvent être faites à chaque fois que le système change d'état afin de garder les postes rouges couverts et de repositionner les ambulances disponibles localisées ailleurs que sur des postes d'attente. Les véhicules sont déplacés par collisions lors des relocalisations.

PARAMETRES

NBROUGE : nombre de postes de priorité rouge

NBNOIR : nombre de postes de priorité rouge et noire.

NOTE : les localisations des postes d'attente sont définies dans le tableau TPOSIT. Les postes prioritaires correspondent aux NBROUGE premières entrées de ce tableau. Les autres entrées se rapportent aux postes non-prioritaires.

MAXAMBREP : nombre maximal de collisions par repositionnement.

TMAXPARCOURS : borne sur le temps de parcours de la deuxième ambulance impliquée dans une chaine de collisions. Si le temps de parcours est supérieur à TMAXPARCOURS, alors la première collision est sautée (voir figure 4).

Justification : le saut de la première collision permet d'éviter une gestion trop lourde des relocalisations lorsque plusieurs ambulances deviennent disponibles en un point à des intervalles de temps courts et réguliers (début de quart dans un même garage).

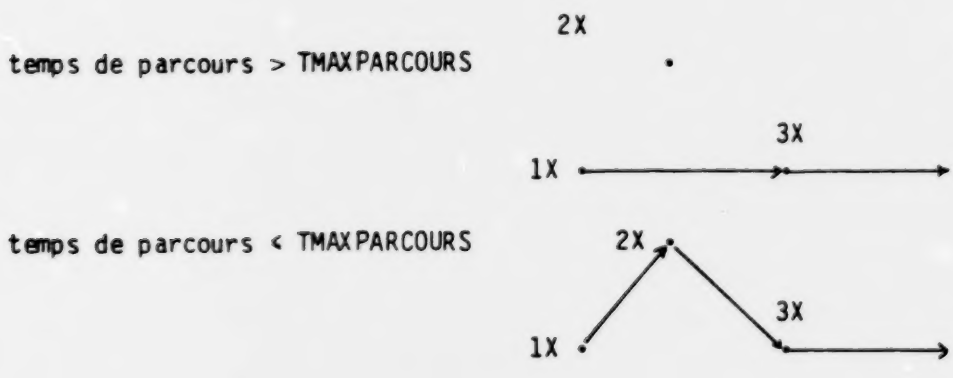

Figure 4

PROCÉDURES (P) ET FONCTION (F)

INITACTIQUE (P) : complète les initialisations pour la tactique Urgences santé no 1 (Lecture des paramètres).

ONLESERT (F) : détermine si un appel est servi immédiatement ou plus tard. Un appel non-prioritaire est placé en file d'attente si le nombre d'ambulances disponibles est inférieur ou égal au nombre de postes de priorité rouge.

RELOCALISER (P) : traite un évènement TRE en deux étapes :

1) identifie toutes les ambulances disponibles localisées ailleurs que sur des postes d'attente. La procédure DEPLACEAMB est appelée pour relocaliser chacune d'elles.
2) identifie tous les postes d'attente de priorité rouge qui sont libres. La procédure COUVREPOSTE est alors activée pour tenter de couvrir chacun d'eux.

DEPLACEAMB (P) : effectue une relocalisation étant donné qu'une ambulance est devenue disponible (suite à un évènement TSF, TDQ, TFP). Le poste libre de priorité supérieure le plus près est identifié. Si un tel poste existe, alors une série de collisions est amorcée avec POSITIONNE (voir simulateur) dans le but de couvrir ce poste vacant.

COUVREPOSTE (P) : effectue une relocalisaton dans le cas où un poste d'attente prioritaire devient libre suite au départ d'une ambulance (causé par un TSS, TFQ ou TDP). Dans un premier temps, on détermine le poste occupé de priorité noire qui est le plus près. Si un tel poste existe, alors dans un deuxième temps, une série de collisions est amorcée afin de couvrir le poste libre à partir de l'ambulance localisée sur le poste de priorité noire.

7.2.2 Urgences santé no 2

Cette famille de tactiques est inspirée de la manière dont Urgences santé gère sa flotte d'ambulances depuis décembre 1985.

PRINCIPE

On dispose d'un ensemble de postes d'attente. Leur nombre ainsi que leurs localisations sont, en principe, déterminés de façon à satisfaire d'une manière optimale certains critères de qualité de service dans un contexte déterministe. Dans un tel contexte, on suppose que chacun des postes est couvert en tout temps. Un poste peut être occupé par plusieurs ambulances. Le but de la tactique est de faire en sorte que chacun des postes soit effectivement couvert en tout temps en gérant adéquatement l'allocation et la relocalisation des ambulances disponibles.

La tactique Urgences santé no 1 se définissait en fonction d'une structure de priorité entre les postes. Le but était de couvrir en tout temps les postes de priorité rouge. Ici, tous les postes doivent être couverts en tout temps. Le but de la tactique ne fait pas intervenir une structure de priorité entre les postes.

Toutefois, une telle structure est nécessaire en pratique afin de gérer l'allocation des véhicules. En particulier, il faut déterminer les postes à couvrir en premier, de même que les postes à couvrir en double en premier (lorsque tous les postes sont occupés). Ainsi, nous allons considérer une structure de priorité à deux niveaux.

1er niveau : priorité lorsqu'il existe des postes découverts. Ici, on privilégie les postes selon la densité d'appels de leur région de couverture et selon leur éloignement géographique.

2^{e} niveau : priorité lorsque tous les postes sont couverts. On prévilégie les postes selon la densité d'appels et selon leur situation stratégique pour effectuer des relocalisations.

Note : Pour simplifier, on exige que les postes d'attente prioritaires de deuxième niveau forme un sous-ensemble des postes prioritaires de premier niveau.

Des relocalisations peuvent être faites :

1) lorsqu'une ambulance devient disponible;
2) lorsqu'un poste devient découvert et qu'il existe un ou plusieurs autres postes couverts en double.

Encore ici, les relocalisations sont faites par collisions.

PARAMÈTRES

NBROUGE : nombre de postes prioritaires de premier niveau.

NBPCD : nombre de postes prioritaires de deuxième niveau.

NOTE : 1) les NBROUGE premières entrées de TPOSIT correspondent aux postes prioritaires de premier niveau, tandis que les NBPCD premières entrées de TPOSIT correspondent aux postes prioritaires de deuxième niveau;

2) NBPCD < NBROUGE toujours.

MAXAMBREP : voir Urgences santé no 1.

TMAXPARCOURS : voir Urgences santé no 1.

PLANCHER : le nombre d'ambulances disponibles au dessous duquel un appel non-prioritaire est placé en file d'attente.

Les procédures (P) et fonction (F)

INITACTIQUE (P) : complète les initialisations pour la tactique Urgences Santé no 2.

ONLESERT (F) : détermine si un appel est servi immédiatement ou plus tard. Si le nombre d'ambulances disponibles est inférieur ou égal au plancher, alors les appels non-prioritaires sont placés en file d'attente.

RELOCALISER (P) : effectue le traitement d'un évènement TRE (en deux étapes):

1) elle identifie toutes les ambulances disponibles localisées ailleurs que sur des postes. La procédure DEPLACEAMB est appelée pour relocaliser chacune d'elles.
2) elle identifie tous le postes libres et amorce, si possible pour chacun d'eux, une relocalisation pour les couvrir.

DEPLACEAMB (P) : effectue une relocalisation étant donné qu'une ambulance disponible est localisée ailleurs que sur un poste d'attente. S'il existe des postes découverts, alors on identifie le poste prioritaire de premier niveau situé le plus près. Si tous les postes sont occupés, alors on identifie le poste prioritaire de deuxième niveau situé le plus près. Une série de collisions est amorcée avec POSITIONNE afin de couvrir le poste identifié.

COUVREPOSTE (P) : tente d'amorcer une relocalisation afin de couvrir un poste devenu découvert. Dans un premier temps, on détermine s'il existe des postes couverts par plus d'une ambulance. Si oui, alors dans un deuxième temps, on identifie celui qui est le plus près du poste à couvrir. Dans un troisième temps, une série de collisions est amorcée.

7.2.3 Garage

On dispose d'un ensemble de postes d'attente. Lorsqu'une ambulance est disponible, elle est localisée sur un de ces postes. Chacune des ambulances est assignée à un poste d'attente pour toute la durée de la simulation. C'est un contexte garage où aucune relocalisation n'est permise. Chacun des postes (garage) peut se voir assigner une ou plusieurs ambulances.

Paramètres

TYPEGARAGE : type de tactique garage. Deux valeurs sont possibles :

0 : le retour en disponibilité des ambulances se fait au C.H.

1 : le retour en disponibilité se fait au garage (contexte de l'hypercube, voir [15]).

FGARAGE : fichier indiquant pour chaque ambulance, le numéro du poste d'attente auquel elle est assignée.

PLANCHER : nombre d'ambulances disponibles au dessous duquel un appel non-prioritaire est placé en file d'attente.

Procédures (P) et fonction (F)

INITACTIQUE (P) : complète les initialisations.

ONLESERT (F) : détermine si un appel est servi immédiatement ou plus tard. Si le nombre d'ambulances disponibles est inférieur ou égal au plancher, alors les appels non-prioritaires sont placés en file d'attente.

RELOCALISER (P) : relocalise les ambulances devenues disponibles en les envoyant directement à leur garage respectif.

Les modules que nous avons présentés jusqu'ici permettent de simuler un système d'urgence avec une politique de fonctionnement donnée. Il est ainsi possible de faire une expérience de simulation durant laquelle des observations sur les appels sont recueillies. Toutefois, nous n'avons pas encore vu comment analyser ces résultats. Ceci fait l'objet du prochain chapitre.

CHAPITRE 8
LE MODULE D'ANALYSE DES RÉSULTATS

8.1 Généralités

Supposons que l'on dispose d'un modèle de simulation (pour un système d'urgence) construit à partir de données réelles. Ce modèle se veut un reflet de la réalité. Nous voulons entre autres nous en servir afin d'évaluer des tactiques de déploiement des véhicules sans les essayer dans la réalité. À cette fin, le modèle doit constituer une bonne approximation du système réel. Nous allons maintenant préciser la nature de cette approximation.

Soit T, une performance du système modélisé (le temps de déplacement pour un appel, par exemple). T est une variable aléatoire. De plus, soit $t_1, t_2, \ldots, t_n$ un ensemble de réalisations de la variable aléatoire T obtenu par une expérience de simulation. Alors deux problèmes se posent.

1) Nous voulons estimer $\theta = E(T)$ (l'espérance du temps de déplacement) avec un certain niveau de fiabilité statistique. En particulier, nous voulons obtenir un intervalle de confiance pour θ. Pour cela, nous avons besoin

   i) d'un bon estimateur de θ ($\bar{T}$)

   ii) d'un bon estimateur de la variance de $\bar{T}$ ($V(\bar{T})$).

2) En supposant que nous pouvons bien estimer θ, nous devons extrapoler les résultats de cette estimation à la réalité. En effet, à partir d'un échantillon obtenu d'une population A (le modèle), nous devons tirer des conclusions sur une population B (le système réel).

Selon Bratley et al. [2, p. 27], il n'y a pas ou peu de fondements statistiques pour le problème 2. Bien sûr, dans le cas d'un modèle dont les règles de décision sont les mêmes que celles d'un système réel, on peut

utiliser certaines techniques de validation (voir chapitre 10). Toutefois, dans le cas d'une tactique de déploiement qui n'a jamais été opérationnelle, il n'y a rien qui nous assure de la validité de l'extrapolation des résultats du modèle à la réalité. La solution à ce problème est en grande partie subjective. Dans ce qui suit, nous discuterons plutôt des techniques utilisées pour résoudre le problème (1).

Une expérience de simulation produit des observations qui, la plupart du temps, sont corrélées. Ainsi, les méthodes classiques d'inférence statistique ne peuvent être employées, car elles présupposent l'indépendance des observations. De plus, elles utilisent l'hypothèse de normalité ou du moins, une théorie asymptotique de la normalité. Ce qui n'est pas toujours vérifié non plus dans des résultats de simulation.

Par conséquent, des méthodes d'analyse statistique propres à la simulation doivent être considérées. Notre propos, ici, n'est pas d'en faire la revue. Le lecteur intéressé à ce sujet peut consulter [2, 3, 4, 5] où des méthodes différentes sont présentées selon le type de simulation (horizon fini ou horizon infini). Le traitement pour le cas horizon infini est plus difficile.

Autant en contexte horizon fini qu'infini, on est intéressé à construire des intervalles de confiance pour les mesures. Il est donc nécessaire d'avoir des méthodes permettant de construire de tels intervalles pour une taille d'échantillon fixée. De plus, il est nettement souhaitable de disposer de procédures itératives afin de déterminer la longueur d'une simulation dans le but d'atteindre un niveau de confiance et une précision donnés à priori sur les intervalles. Nous discutons à présent des procédures que nous avons implantées pour chacun des deux types de simulation.

8.2 Contexte horizon fini

Supposons que l'on fait n répétitions indépendantes d'une simulation en contexte horizon fini. La période considérée peut être, par exemple, une journée d'opération débutant à 06 hres. et se terminant vingt-quatre heures plus tard. Pour chaque répétition, les conditions intiales sont générées à partir d'une même loi de probabilité. De plus, l'indépendance entre chacune de ces répétitions est obtenue en utilisant des nombres aléatoires différents pour chacune d'elles.

Il est alors facile de construire un intervalle de confiance de confiance dans ce contexte. Supposons que nous désirions estimer $E(T)=\Theta$, l'espérance du temps de déplacement. Pour chaque répétition i, on définit

$$T_i = \frac{\sum_{j=1}^{n_i} T_{ij}}{n_i}$$

où T_{ij} : j^e temps de déplacement dans la $i^{ème}$ répétition;

n_i : nombre de temps de déplacement dans la $i^{ème}$ répétition.

On dispose donc de l'échantillon $T_1, T_2, \ldots, T_n$. Les T_i sont des variables aléatoires indépendantes et issues d'une même loi de probabilité (voir [3, p. 288]). Si les T_i sont issues d'une loi normale, alors un intervalle de confiance pour Θ de niveau $1-\alpha$ est donné par

$$\bar{T} \pm t_{n-1;\, 1-\alpha/2} \sqrt{\frac{S^2(n)}{n}} \qquad (1)$$

$$\text{où } \bar{T} = \frac{\sum_{i=1}^{n} T_i}{n},$$

$$S^2 = \frac{\sum_{i=1}^{n} (T_i - \bar{T})^2}{n-1}.$$

et $t_{n-1;\, 1-\alpha/2}$ est tel que $P[\text{Student } n-1 < t_{n-1;\, 1-\alpha/2}] = 1-\alpha/2$.

Par contre, si les T_i ne sont pas issues d'une loi normale, ce qui est souvent le cas en pratique, alors l'intervalle (1) peut quand même être considéré. En effet, à l'aide du théorème de la limite centrale, on montre que, quel que soit la fonction de répartition des T_i, l'intervalle est approximatif et son niveau de confiance tend vers le niveau désiré lorsque n croît ([2, p. 289]). Dans ce cas, on doit s'assurer que $\bar{T}$ est approximativement normal ([2, p. 79]).

Nous définissons la précision d'un intervalle de confiance par sa demi-longueur. Une procédure séquentielle est utilisée afin de déterminer la taille de la simulation en fonction d'une précision et d'un niveau de confiance donnés. Nous avons implanté celle qui est décrite dans [3, p. 292].

Elle se résume comme suit :

0) obtenir n_0 répétitions; $n=n_0$.

1) calculer $\bar{T}(n)$, $S^2(n)$

et $\delta(n,\alpha) = t_{n-1;\ 1-\alpha/2} \sqrt{\frac{S^2(n)}{n}}$

2) si $\frac{\delta(n,\alpha)}{\bar{T}(n)} < \gamma$, alors

utiliser $I(\alpha,\gamma) = [\bar{T}(n)-\delta(n,\alpha),\ \bar{T}(n)+\delta(n,\alpha)]$ comme intervalle de confiance, sinon, aller à 1.

NOTE : γ est la précision relative désirée sur l'intervalle et α est le niveau de confiance.

L'exécution d'une simulation horizon fini se fait en utilisant la procédure SIMHRZFINIE du module d'analyse statistique. Les performances qui peuvent être estimées sont :

- l'espérance du temps de déplacement;
- la proportion d'appels pour lesquels le temps de déplacement est inférieur à cinq minutes.

Dans le cas où l'on estime les deux performances, l'inégalité de Bonferoni est utilisée afin de fixer la valeur critique $t_{n-1;\ 1-\alpha/2}$ (voir [3, p. 308]). La procédure SIMHRZFINIE utilise les paramètres suivants :

MAXJOURS : nombre maximal de journées à simuler.
PRECREL : précision relative sur les mesures.
NIVEAU : niveau de confiance global sur les intervalles de confiance.

La figure 5 montre un exemple d'une analyse d'une simulation horizon fini.

Le problème des conditions intiales

Les conditions initiales définissent l'état du système au début de la simulation. Elles sont issues d'une même loi de probabilité pour chacune des répétitions. Il suffit de générer le nombre de véhicules occupés, et pour chacun d'eux, le temps auquel il redeviendra disponible.

8.3 Contexte horizon infini

L'analyse statistique des résultats d'une simulation horizon infini est plus difficile que dans le contexte horizon fini. Parmi les techniques existantes pour la construction d'intervalles de confiance, nous avons implanté la méthode "BATCH MEANS" ([3, p. 295]). Nous présentons d'abord la méthode pour une simulation de taille fixe p*, en utilisant toujours l'exemple avec le temps de déplacement.

---

*NOTE : Dans un contexte horizon infini, on ne fait qu'une seule longue simulation, plutôt que plusieurs expériences indépendantes ([3, p. 307]).

STATISTIQUES SUR LES PERFORMANCES

NOMBRE DE JOURNÉES SIMULÉES : 100

TEMPS DE DÉPLACEMENT :

MOYENNE	: 5.2555
VARIANCE	: 0.0659
INTERVALLE	: [4.7524 , 5.7586]
PRÉCISION RELATIVE	: 0.0957

PROB[TEMPS < = 5 MINUTES] :

MOYENNE	: 0.9213
VARIANCE	: 0.0015
INTERVALLE	: [0.8463 , 0.9963]
PRÉCISION RELATIVE	: 0.0814

NIVEAU DE CONFIANCE POUR UN INTERVALLE : 95%

Figure 5 : Analyse d'une simulation à horizon fini

Soit Θ, le paramètre à estimer, et $T_1, T_2, \ldots, T_p$ les p temps de déplacement obtenus durant la simulation après qu'on ait éliminé les premiers appels de la période de transition. Supposons que l'équilibre est atteint. On construit alors m paquets ("batches") comportant ℓ observations ($m\ell=p$). On définit pour chacun des paquets i, i=1, ..., m, le temps moyen de déplacement par

$$\bar{T}_i(\ell) = \frac{\sum_{j=1}^{\ell} T_{(i-1)\ell+j}}{\ell} .$$

On peut alors estimer Θ sans biais par

$$\bar{\bar{T}}(m,\ell) = \frac{\sum_{i=1}^{m} \bar{T}_i(\ell)}{m} .$$

Puisqu'on est dans une situation à l'équilibre, alors $\bar{T}_1(\ell)$, $\bar{T}_2, (\ell), \ldots, \bar{T}_m(\ell)$ sont des variables aléatoires identiquement distribuées de moyenne Θ et de variance σ^2 inconnue. De plus, si on choisit ℓ assez grand, ces variables seront approximativement indépendantes à cause de l'effet de corrélation qui ne se fera sentir qu'au début et à la fin de chacun des paquets (voir [3]). Si m est assez grand, alors par le théorème de la limite centrale, $\bar{\bar{T}}(m,\ell)$ est approximativement $N(\Theta, \frac{\sigma^2}{n})$ et par conséquent, un intervalle de confiance approximatif de niveau $1-\alpha$ pour Θ est donné par

$$\bar{\bar{T}}(m,\ell) \pm t_{n-1;\alpha/2} \sqrt{\frac{S^2[\bar{T}_i(\ell)]}{m}}$$

$$\text{où} \quad S^2[\bar{T}_i(\ell)] = \frac{\sum_{i=1}^{m} (\bar{T}_i(\ell) - \bar{\bar{T}}(m,\ell))^2}{m-1} .$$

Afin de déterminer la longueur d'une simulation horizon infini, nous avons implanté une procédure séquentielle dûe à Law et Carson [18]. C'est

une méthode empirique qui permet de réduire la corrélation entre les paquets en augmentant graduellement leur taille et, du même coup, d'atteindre un degré de précision spécifié a priori.

L'exécution d'une simulation à horizon infini se fait en utilisant la procédure SIMSTEADYSTE du module d'analyse statistique. Les performances pouvant être estimées sont les mêmes que dans le contexte horizon fini. Les paramètres de la procédure sont les suivants :

PRECREL : précision relative désirée sur les intervalles de confiance.
NIVEAU : niveau de confiance pour les intervalles.
NBAPPRESIM : nombre d'appels à éliminer du début de la simulation.
NBSURPLUS : nombre d'appels à rajouter lorsque la précision et l'indépendance approximative entre les "batches" sont obtenues.
NBAPPMIN : nombre d'appels minimal désiré.

En sortie, la procédure fournit :

- pour chaque performance :
  - la moyenne;
  - la variance;
  - l'intervalle de confiance;
  - la précision relative et absolue de l'intervalle;
  - le coefficient de corrélation échantillonnale entre les paquets.
- les résultats des itérations de la procédure de Law et Carson.
- les statistiques suivants :
  - la probabilité de chacun des états du système (un état est défini par le nombre de véhicules disponibles);
  - le temps à la fin de la simulation;
  - le nombre d'appels simulés.

La figure 6 illustre l'analyse d'une simulation à horizon infini.

```
STATISTIQUES :

TEMPS DE REPONSE :

MOYENNE                 :      3.7509
VARIANCE                :      0.0207
I.C.                    : [    3.7062 ,      3.7955 ]
PRECISION ABSOLUE       :      0.0446
PRECISION REL.          :      0.0119
COEFF. CORR. -LAG1      :      0.0625

PROB[ TEMPS <= 5 MINUTES ] :

MOYENNE                 :      0.7638
VARIANCE                :      0.0003
I.C.                    : [    0.7581 ,      0.7694 ]
PRECISION ABSOLUE       :      0.0057
PRECISION REL.          :      0.0074
COEFF. CORR. -LAG1      :      0.0134

NIVEAU DE CONFIANCE POUR UN INTERVALLE : 95 %.


NOMBRE D'APPELS SUPPLEMENTAIRES SIMULES                 :      20000
NOMBRE TOTAL D'APPELS SIMULES                           :      39200

PROBABILITES A L'ETAT D'EQUILIBRE :

P[ N= 0 ]=     5.1716522704512E-007
P[ N= 1 ]=     6.1333214120037E-005
P[ N= 2 ]=     4.5097111267310E-004
P[ N= 3 ]=     1.9183180493637E-003
P[ N= 4 ]=     6.7508348814041E-003
P[ N= 5 ]=     2.3071129242249E-002
P[ N= 6 ]=     6.7544930415413E-002
P[ N= 7 ]=     1.5536394684167E-001
P[ N= 8 ]=     2.6483567224500E-001
P[ N= 9 ]=     3.0021674472012E-001
P[ N=10 ]=     1.7978560204652E-001

NOMBRE D'APPELS RETRANCHES DU DEBUT DE LA SIMULATION :            0
TEMPS AU DEBUT DE LA SIMULATION ( APRES PRE-SIMUL. ) :        0.0000
TEMPS A LA FIN DE LA SIMULATION                      :   747762.8648
```

Figure 6 : Analyse d'une simulation à horizon infini

(10 ambulances)

Le problème du biais initial

Dans une simulation à horizon infini, les performances à estimer sont définies comme des limites lorsque le temps tend vers l'infini. On parle alors d'un système à l'équilibre. À cet effet, les performances ne dépendent plus des conditions initiales du système. Toutefois, dans une simulation, les premières observations ne sont pas représentatives de cet état d'équilibre. C'est ce qu'on appelle communément le problème du biais initial.

Afin de résoudre ce problème, il n'existe pas encore de méthode rigoureuse ayant fait ses preuves. Il existe plusieurs règles heuristiques dont la plupart se sont averées inéfficaces [19]. D'autre part, Schruben [20,2] propose un test pour le biais initial, mais sa méthode n'a pas encore été testée suffisamment.

La méthode que nous utilisons est heuristique et inspirée de la technique de Schruben. Elle permet d'examiner, d'une manière descriptive, s'il y a un biais dans une séquence d'observations obtenues d'une expérience de simulation. On considère un échantillon de taille n. Dans un premier temps, un intervalle de confiance est construit avec la deuxième moitié de l'échantillon. Dans un deuxième temps, une suite d'intervalles est construite en incluant graduellement les observations de la première motié de l'échantillon. On observe finalement la séquence d'intervalles obtenues (ex. voir figure 7). Si les derniers intervalles "bougent" beaucoup par rapport aux premiers, on peut alors présumer la présence d'un biais initial.

En plus de cette méthode, il est d'usage d'examiner le graphique des performances cumulatives (voir figure 8).

ANALYSE DES INTERVALLES DE CONFIANCE

NOMBRE D'OBSERVATIONS	18000	
MOYENNE	3 7534	
VARIANCE	0 0193	
I C (40 BATCHS)	[3 7315 , 3 7754]	* M *
LONGUEUR DES BATCHS	150	
NOMBRE D'OBSERVATIONS	21000	
MOYENNE	3 7456	
VARIANCE	0 0148	
I C (40 BATCHS)	[3 7264 , 3 7648]	* M *
LONGUEUR DES BATCHS	524	
NOMBRE D'OBSERVATIONS	24000	
MOYENNE	3 7469	
VARIANCE	0 0148	
I C (40 BATCHS)	[3 7276 , 3 7661]	* M *
LONGUEUR DES BATCHS	600	
NOMBRE D'OBSERVATIONS	27000	
MOYENNE	3 7342	
VARIANCE	0 0126	
I C (40 BATCHS)	[3 7165 , 3 7520]	* M *
LONGUEUR DES BATCHS	674	
NOMBRE D'OBSERVATIONS	30000	
MOYENNE	3 7377	
VARIANCE	0 0112	
I C (40 BATCHS)	[3 7209 , 3 7544]	* M *
LONGUEUR DES BATCHS	750	
NOMBRE D'OBSERVATIONS	33000	
MOYENNE	3 7415	
VARIANCE	0 0100	
I C (40 BATCHS)	[3 7257 , 3 7573]	* M *
LONGUEUR DES BATCHS	824	
NOMBRE D'OBSERVATIONS	36000	
MOYENNE	3 7400	
VARIANCE	0 0099	
I C (40 BATCHS)	[3 7243 , 3 7557]	* M *
LONGUEUR DES BATCHS	900	

Figure 7 : Détection du biais initial

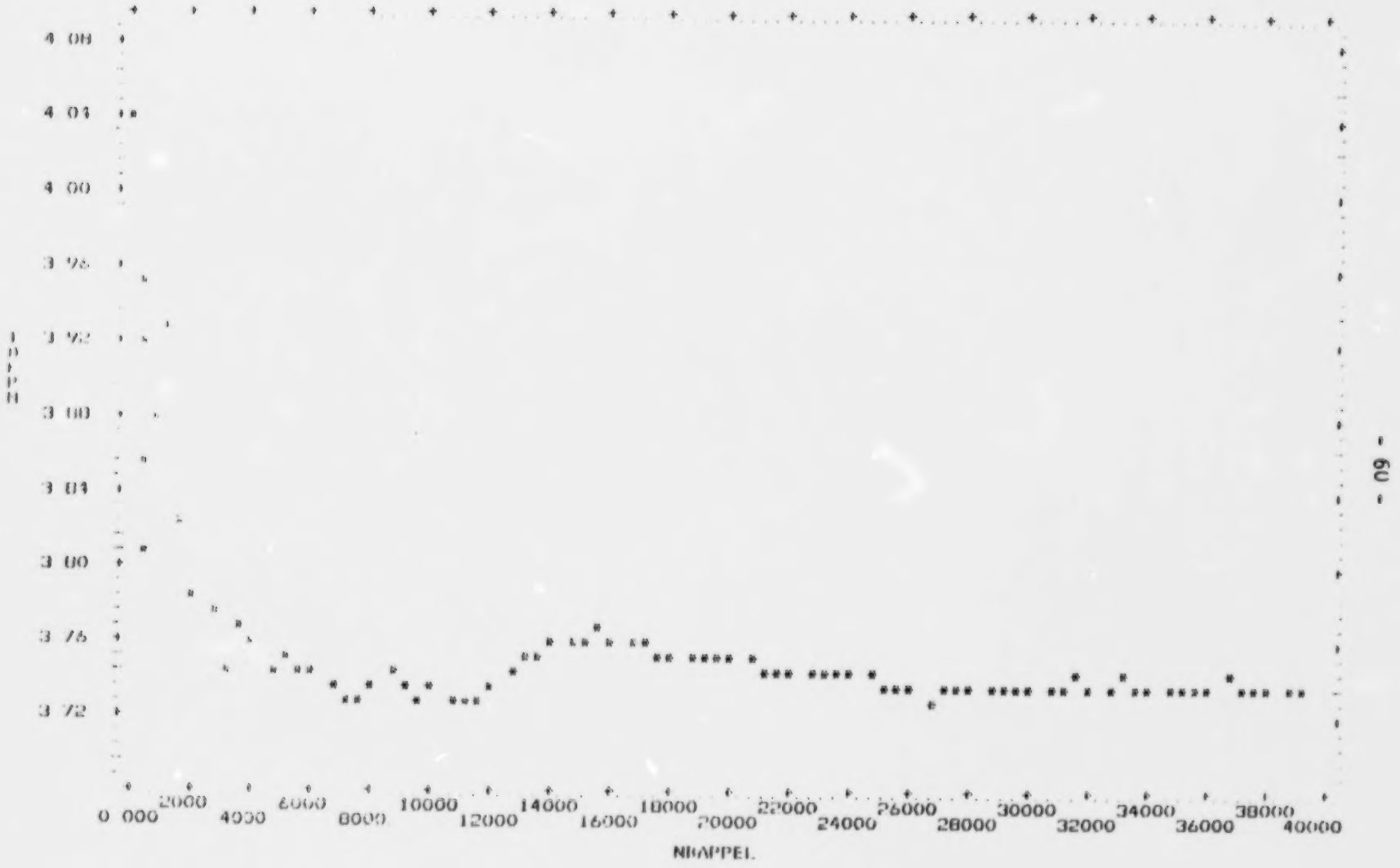

Figure 8 : Temps de déplacement moyen cumulatif

Il est à remarquer qu'avec la présente méthode, il n'est pas possible de détecter la présence d'un biais durant le déroulement de la simulation. En conséquence, il est préférable de simuler un nombre d'appels supplémentaires dans la procédure SIMSTEADYSTE (paramètre NBSURPLUS).

8.4 Autres procédures statistiques

Nous avons également implanté une procédure permettant d'estimer par intervalles n charges de travail (voir figure 9). La méthode est décrite dans [14, p. 90]. De plus, des programmes permettent de calculer différentes statistiques descriptives à partir du fichier de résultats, et ce, pour les contextes horizon fini et infini (voir figure 10).

CHARGES DE TRAVAIL

```
UNITE[ 1]
MOYENNE                          =    0.1851
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1783 , 0.1919 ]
PRECISION DE L'INTERVALLE        =    0.0068

UNITE[ 2]
MOYENNE                          =    0.1877
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1809 , 0.1944 ]
PRECISION DE L'INTERVALLE        =    0.0067

UNITE[ 3]
MOYENNE                          =    0.2128
VARIANCE                         =    0.0003
INTERVALLE DE CONFIANCE          =  [ 0.2050 , 0.2206 ]
PRECISION DE L'INTERVALLE        =    0.0078

UNITE[ 4]
MOYENNE                          =    0.1641
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1577 , 0.1706 ]
PRECISION DE L'INTERVALLE        =    0.0065

UNITE[ 5]
MOYENNE                          =    0.0977
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.0916 , 0.1037 ]
PRECISION DE L'INTERVALLE        =    0.0061

UNITE[ 6]
MOYENNE                          =    0.1319
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1263 , 0.1376 ]
PRECISION DE L'INTERVALLE        =    0.0057

UNITE[ 7]
MOYENNE                          =    0.1979
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1909 , 0.2049 ]
PRECISION DE L'INTERVALLE        =    0.0070

UNITE[ 8]
MOYENNE                          =    0.2059
VARIANCE                         =    0.0003
INTERVALLE DE CONFIANCE          =  [ 0.1984 , 0.2133 ]
PRECISION DE L'INTERVALLE        =    0.0075

UNITE[ 9]
MOYENNE                          =    0.1521
VARIANCE                         =    0.0003
INTERVALLE DE CONFIANCE          =  [ 0.1443 , 0.1599 ]
PRECISION DE L'INTERVALLE        =    0.0078

UNITE[10]
MOYENNE                          =    0.2045
VARIANCE                         =    0.0002
INTERVALLE DE CONFIANCE          =  [ 0.1975 , 0.2115 ]
PRECISION DE L'INTERVALLE        =    0.0070
```

Figure 9 : Estimation par intervalles des charges de travail

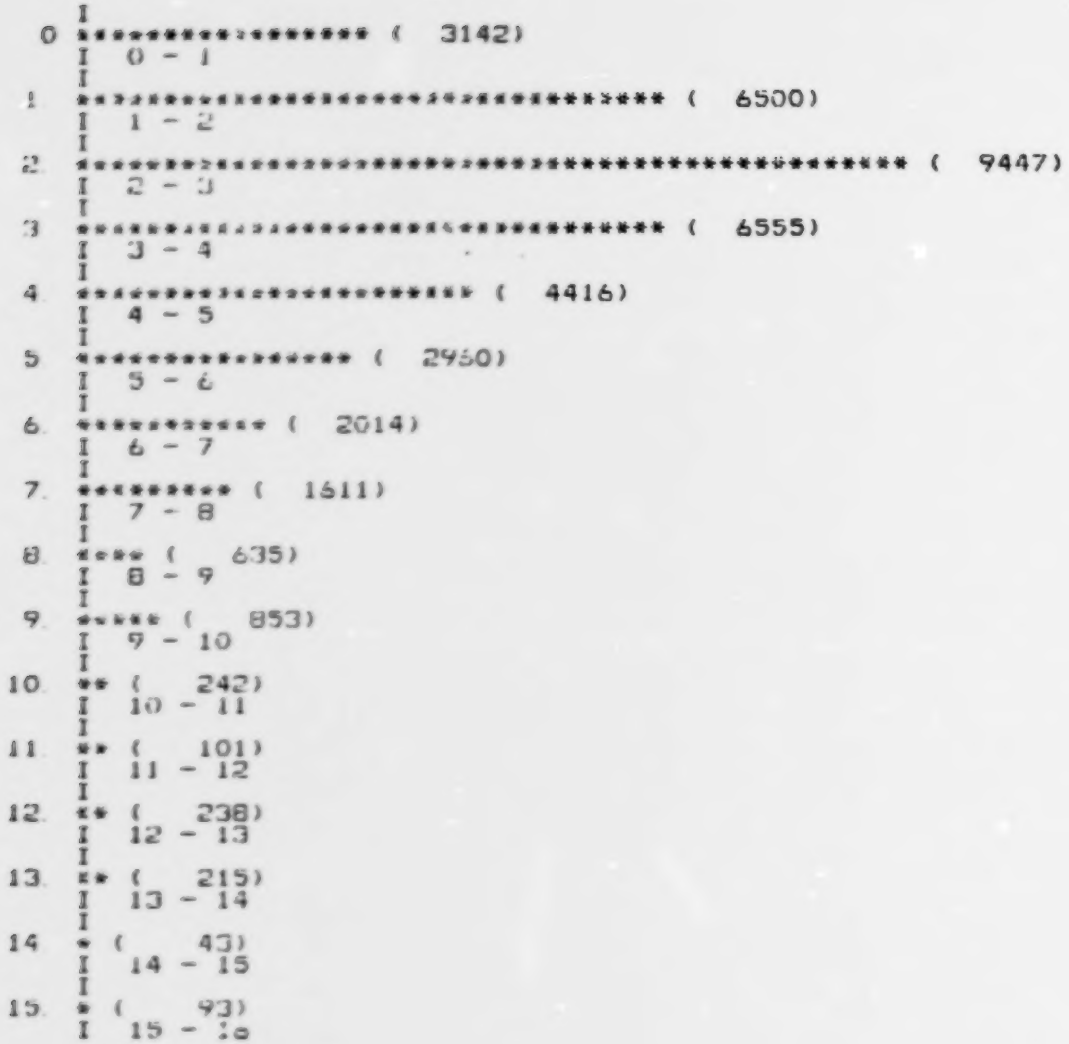

Figure 10 : Histogramme pour le temps de déplacement

CHAPITRE 9
LE SYSTÈME D'INTERFACE

9.1 Généralités

Le modèle de simulation que nous avons présenté jusqu'à maintenant comporte plusieurs modules. Lors d'une expérience de simulation, il est nécessaire de faire exécuter différents programmes qui sont répartis entre ces modules (voir le manuel de l'usager [21]). Un système d'interface est alors indispensable pour :

- définir les paramètres d'une simulation;
- assurer un lien cohérent entre les différents modules;
- emmagasiner les informations relatives à chaque appel, de même que l'état du système à la fin d'une simulation.

Un système d'interface comporte principalement deux éléments :

1) une structure de fichiers efficace;
2) une procédure interactive permettant de créer, de consulter et de mettre à jour automatiquement les fichiers de paramètres afin d'éviter les erreurs dues à la manipulation de ces fichiers par l'usager.

NOTE : En ce qui concerne le point (2), nous avons, pour le moment, développé une procédure qui permet de ne sélectionner que les paramètres propres au choix de la tactique (et donc pas pour la sélection des paramètres globaux). Une procédure interactive complète constituerait normalement un élément d'un système de gestion du modèle. Le but d'un tel système serait de permettre à tout usager d'effectuer facilement des expériences de simulation du début à la fin. Il devrait entre autres offrir des possibilités graphiques pour l'analyse des résultats et un système de gestion de base de données pour emmagasiner les différentes informations concernant les données du modèle, les paramètres et les résultats des expériences de simulation.

Dans ce qui suit, nous allons principalement parler de la structure de fichiers utilisée.

9.2 Description de la structure de fichiers

9.2.1 La structure globale

Les différents programmes du modèle doivent être alimentés en entrée. Des fichiers d'entrée sont nécessaires pour :

- le générateur des demandes;
- le programme interactif de sélection des tactiques;
- le simulateur;
- le module de décision;
- les procédures statistiques.

De plus, des fichiers de sortie sont créés par chacun de ces programmes (sauf les procédures statistiques) afin d'être utilisés en entrée pour d'autres programmes. La figure 11 montre les relations entre chacun des modules. C'est à partir de ce canevas que la structure de fichiers a été développée.

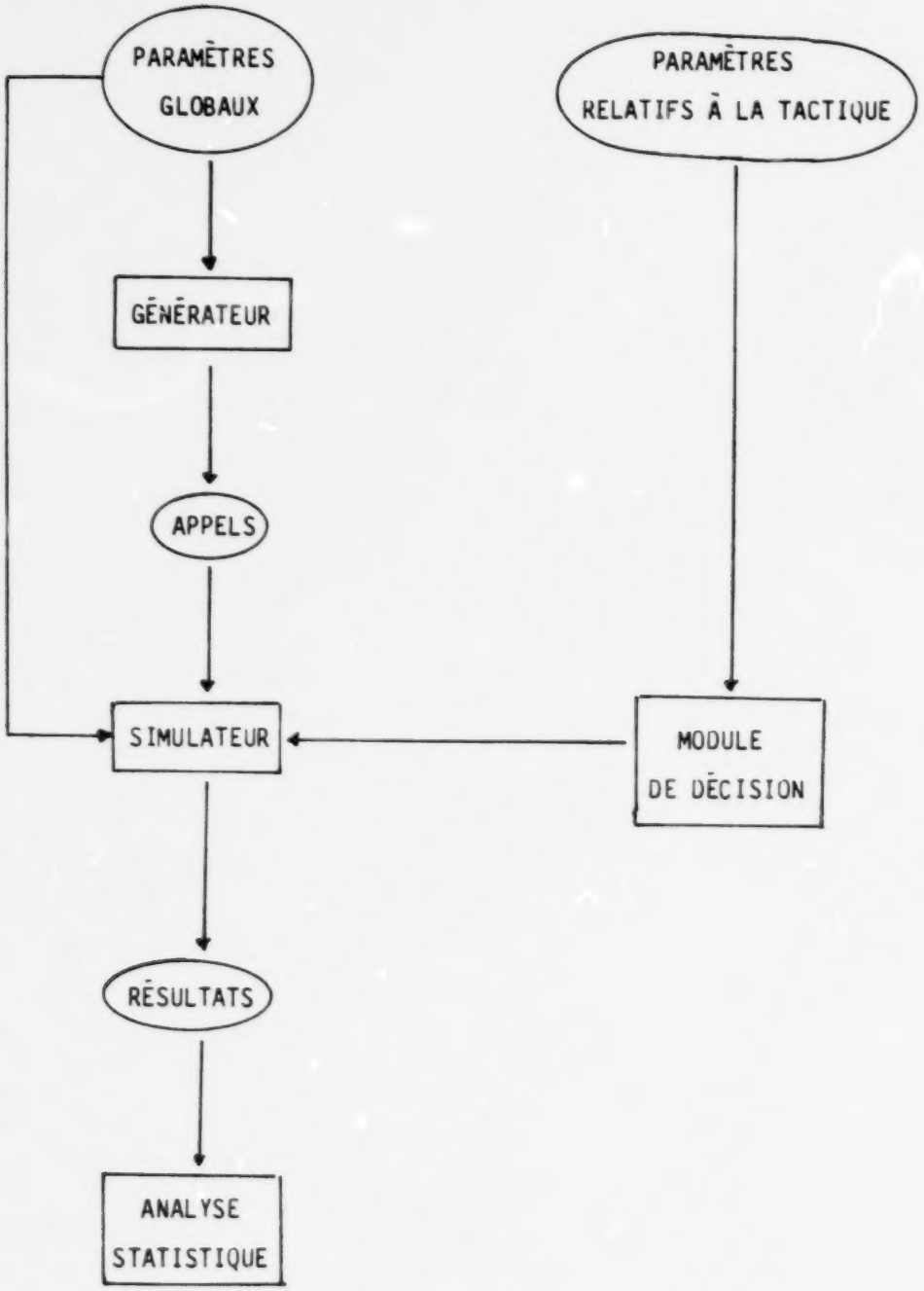

Figure 11 : LIENS ENTRE LES DIFFÉRENTS MODULES

Afin d'éviter d'alourdir inutilement l'utilisation du modèle, l'usager devrait pouvoir spécifier les paramètres d'une simulation en créant un nombre minimal de fichiers. Dans une étude de simulation, on utilise souvent un même fichier d'appels pour simuler plusieurs tactiques. Dans ce cas, il est nécessaire que les paramètres globaux et les paramètres propres à chaque tactique soient localisés en un seul endroit. Afin de regrouper ces informations, nous utilisons les fichiers permanents de groupe 2 disponibles sur CYBER ([22]). Au début de toute expérience de simulation, un tel fichier permanent est créé. Ce dernier constitue un fichier de fichiers dont nous allons décrire sommairement chacune des composantes. Le lecteur intéressé par les détails (contenu et format de chaque fichier) devra consulter le manuel de l'usager ([21]).

9.2.2 Description des fichiers de paramètres résidants sur le fichier permanent de base

Fichiers pour la description des paramètres globaux :

FPARAMA : contient les paramètres globaux (voir section 4.3) tels que spécifiés par l'usager. Il sert d'entrée au générateur.

FPARAM : fichier de sortie du générateur. Il sert d'entrée au simulateur.

FAPPEL : fichier de sortie du générateur. Il contient les informations pour chaque appel.

FCOMMENT : fichier de commentaires décrivant chacune des tactiques déjà programmées sur le fichier permanent. Il est créé par la procédure interactive pour la sélection des tactiques (CREFIC2).

Fichiers relatifs à chacune des tactiques

Nous avons développé une procédure interactive afin de faciliter la sélection d'une tactique et des différents paramètres qui s'y rapportent.

Les fichiers relatifs à chacune des tactiques sont créés par CREFIC2 et servent d'entrée au simulateur. Les fichiers pour la première tactique considérée sont catalogués sur u=1, ceux pour la deuxième, sur u=2, et ainsi de suite. Pour chacune des tactiques, les deux fichiers suivants sont toujours présents :

FPOSTE : information sur les postes d'attente;

FLANCE : contient les cartes de contrôle pour faire exécuter le programme de simulation.

Les fichiers propres à chaque tactique sont les suivants :

1) Urgences santé no 1

   RGTRI : fichiers d'arcs triés entre les postes de priorité rouge

   NRTRI : fichiers d'arcs triés entre les postes de priorité rouge et noire

2) Urgences santé no 2

   RGTRI : fichiers d'arcs triés entre les postes d'attente

3) Garage

   FGARAGE : indique à quel poste chacune des ambulances est affectée

La figure 12 représente un fichier permanent pour une expérience de simulation où les tactiques 1, 2 et 3 y sont comparées. Ces dernières ont été créées dans l'ordre. La figure 13 montre en détail les relations entre les fichiers et les différents modules du système.

NOTE : Le fichier de résultat pour chaque simulation n'est pas catalogué sur le fichier permanent car il est généralement trop gros. Il est plutôt catalogué sur SN=groupe2.

FPARAMA	
FQUART	
FPARAM	
FCOMMENT	
FAPPEL	
FPOSTE	u=1
RGTRI	u=1
NRTRI	u=1
FLANCE	u=1
FPOSTE	u=2
RGTRI	u=2
FLANCE	u=2
FPOSTE	u=3
FGARAGE	u=3
FLANCE	u=3

Figure 12 : Fichier permanent pour une expérience de simulation

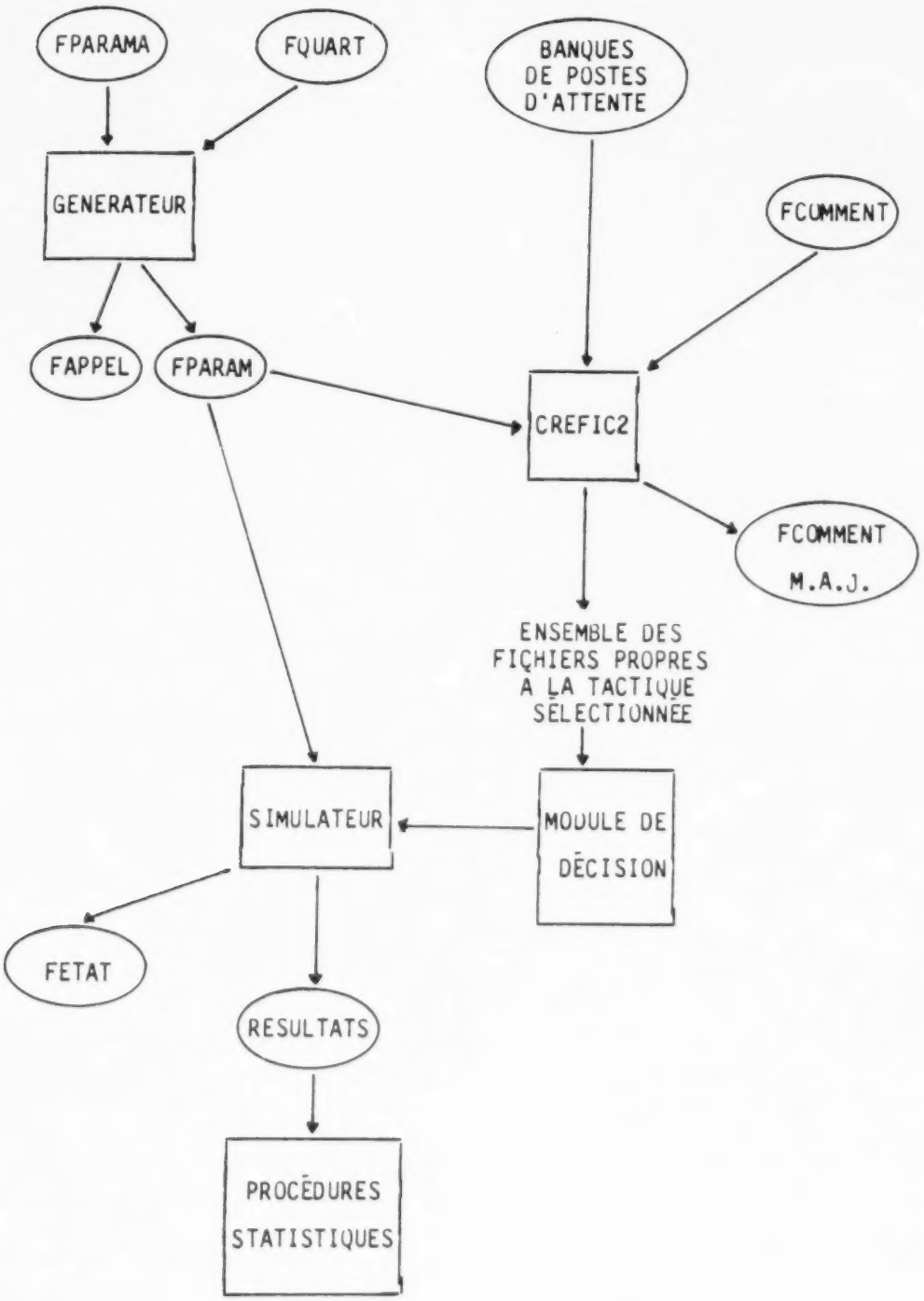

Figure 13 : Liens entre les fichiers et les modules du système

9.2.3 La banque de postes d'attente

Lors de la sélection d'une tactique à l'aide du programme CREFIC2, l'usager détermine le type de tactique et sélectionne ensuite les paramètres qui lui sont propres. Un des paramètres est la liste de noeuds identifiant chacun des postes d'attente. Pour une tactique utilisant le même ensemble de postes, il peut y avoir plusieurs variantes. Par exemple, on peut vouloir évaluer l'effet de différents taux d'appels pour une solution donnée. Ainsi, une banque dans laquelle sont emmagasinés chacun des ensembles de postes d'attente permet de générer facilement plusieurs variantes pour une solution donnée, car l'information est centralisée. La banque de postes d'attente est utilisée en entrée avec le programme CREFIC2. Pour plus de détails, voir [21].

CHAPITRE 10
LE DÉVELOPPEMENT DU MODÈLE

10.1 Le cas particulier : Urgences santé

Dans les chapitres précédents, nous avons tenter d'adopter une approche générale dans la présentation du modèle pour montrer qu'il peut être adapté à une variété de contextes (pompiers, police, ambulances) et de villes. Malgré cela, le modèle présenté est orienté en fonction de son application au système d'ambulances de Montréal. C'est pour cette raison que certains aspects abordés étaient particuliers à ce système, notamment en ce qui concerne la définition des tactiques au chapitre 7.

Dans le présent chapitre, nous allons compléter la description de certaines particularités du cas d'Urgences Santé. Afin de cerner ces éléments, il est nécessaire de décrire les différentes étapes dans le développement du modèle :

1) formulation du problème et plan de l'étude;
2) collecte de données;
3) définition du modèle, formulation des hypothèses, modélisation et validation;
4) programmation du modèle - vérification ("debugging")
5) validation du modèle programmé sur des données réelles et ajustements.

Dans [1] on discute en détail du point (1). Dans les chapitres précédents, nous avons abordé la définition du modèle et sa programmation. Les aspects particuliers que nous allons maintenant traiter sont donc la collecte de données et la validation du modèle.

10.2 La collecte de données

Au chapitre 3, nous avons discuté des informations nécessaires. Nous allons décrire comment celles-ci ont été obtenues.

10.2.1 Les données de base

Nous avons adopté l'approche "réseau" pour représenter le territoire (voir chapitre 3) et pour construire la table des distances. Nous avons procédé en cinq étapes :

1) création d'un réseau agrégé validé

Il permet de construire la table de distances entre les centroïdes de secteurs de recensement issues des données de Statistiques Canada (1981), (nombre de noeuds : 5191, nombre d'arcs : 7707).

Les principales opérations sont l'extraction des données [10] et la construction du réseau [8, p. 10 et 81].

2) division du territoire en secteurs

Chaque secteur est caractérisé par un polygone et par un centroïde. Des coordonnées (x,y) définissent ce centroïde et les sommets du polygone. 484 secteurs issus des données de Statistiques Canada ont été utilisés [10]. La structure de ces données est décrite dans [12].

NOTE : les données originales de Statistiques Canada comportaient 495 secteurs de recensement. Certains ont été fusionnés, car ils étaient non-connexes ou de forme irrégulière. De plus, certains centroïdes ont été déplacés car ils étaient situés dans des zones de faible densité de population.

3) création d'un réseau dilaté

Nous considérons trois (3) zones de congestion. La longueur des arcs d'une zone est multipliée par un facteur de dilatation. Pour les zones plus congestionnées, ce facteur est plus élevé.

Nous n'avons pu avoir de données d'Urgences santé sur la calibration des vitesses. Nous avons utilisé les zones et les facteurs de dilatation définis dans le cadre du projet sur le transport des handicapés à Montréal [8].

4) construction de la table des distances

La distance entre deux centroïdes de secteurs de recensement est obtenue en calculant le plus court chemin qui les relie dans le réseau créé en (3). De plus, pour chaque secteur nous avons calculé une distance intra-secteur de la façon suivante :

$$d = \frac{1}{2}\sqrt{\text{aire du secteur}}$$

Références : [8, p. 25] pour les distances inter-secteur
[15, p. 135] pour les distances intra-secteur.

5) construction d'un tableau de prédécesseurs

Les informations fournies par la table des distances ne permettent pas de connaître la position d'un véhicule se déplaçant le long du plus court chemin d'un noeud i à un noeud j. Toutefois, il est possible de connaître "exactement" cette position si on travaille avec le réseau agrégé. Il suffit d'utiliser la liste des successeurs de i à j obtenus lors du calcul du plus court chemin. Il est cependant impossible pour des raisons d'efficacité de travailler avec ce réseau en simulation.

Nous avons plutôt construit un réseau dont l'ensemble des noeuds est constitué des 484 centroïdes et dont l'ensemble des arcs a été déterminé heuristiquement. On a calculé ensuite les plus courts chemins entre chaque paire de noeuds sur ce réseau "approximatif" pour obtenir un tableau de successeurs de i à j, i,j=1... 484. Ce tableau est utilisé conjointement avec la table des distances crée en (4).

L'ensemble des arcs du réseau approximatif est déterminé par la procédure CRSRES. Le but est d'obtenir un réseau connexe avec un nombre d'arcs restreint de telle sorte que pour toute paire de noeuds, le plus court chemin approximé soit assez semblable au vrai plus court chemin.

Procédure CRSRES :

Elle consiste à prendre chacun des triplets de noeuds et à examiner les trois arcs qui y sont associés.

- Initialisations :

On suppose l'existence d'un réseau complet, chaque arc (a,b) étant de longueur d(a,b) issue de la table des distances. Au départ chaque arc est de statut 0. Il existe trois statuts :

0 : l'arc n'a pas été examiné;

1 : l'arc a été examiné et n'est par rejeté;

2 : l'arc a été examiné et est définitivement rejeté.

- Examen de chaque triplet de noeuds (a,b,c)

Soit (a,b), l'arc le plus long des trois arcs (voir figure 14)

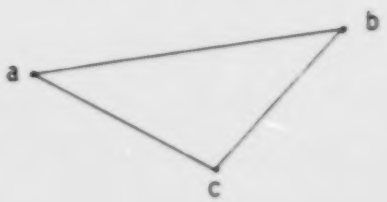

Figure 14 : Examen d'un triplet de noeuds

L'arc (a,b) sera rejeté si

$$\text{facteur} \times d(a,b) > d(a,c) + d(c,b)$$

Justification : facteur est un nombre réel supérieur à 1. Ainsi (a,b) est rejeté s'il est peu ou pas désavantageux de passer par c pour aller de a à b.

Si (a,b) est rejeté alors son statut devient 2, sinon son statut devient 1.

Remarque : Pour obtenir un résultat satisfaisant, on procède par essai-erreur en variant les paramètres du programme. Nous avons finalement utilisé facteur = 1.2.

6) les fonctions de vitesses

Nous avons besoin de deux fonctions de vitesse : une pour les déplacements urgents et l'autre pour les déplacements courants. Il serait intéressant de considérer des fonctions qui tiennent compte de la distance à parcourir. On pourrait de cette façon tenir compte du temps relatif à l'accélération et à la décélération pour les petites distances, de même que de la possibilité d'emprunter des voies rapides pour les plus grandes distances (voir [8]). Malheureusement, nous n'avons pas pu définir de telles fonctions en raison d'absence de données pour calibrer les vitesses sur le réseau. Nous avons supposé les vitesses moyennes suivantes :

- 50 km/h pour les déplacements urgents (suggéré par Urgences santé);
- 27 km/h pour les déplacements courants (selon une étude sur le déplacement des ambulances à Boston [13]).

7) localisation des C.H.

Nous avons considéré les C.H. les plus importants, i.e. ceux ayant reçus plus de cinquante patients pour le mois de juin 1984. Les C.H. ont été localisés sur le réseau agrégé à l'aide des programmes décrits dans [8].

10.2.2 Les variables aléatoires

L'ensemble des variables aléatoires a trait à l'arrivée des appels et au service (voir chapitre 3). Nous avons pu obtenir des données d'Urgences santé pour ces informations. Ces dernières couvrent la période du mois de juin 1984. Dans [14], on décrit en détail l'analyse de données qui en a été faite. Nous énumérons, ici, les informations obtenues pour chaque élément.

Il est à noter que nous n'avons considéré que les appels d'urgences (et donc pas les appels inter-établissement).

1) arrivée des appels

- les taux d'appels

Dans [14], on s'est assuré que l'arrivée des appels, pour l'ensemble du territoire, suit une loi de Poisson sur des périodes de temps où le phénomène est à peu près constant. De plus, on a identifié trois types de journées au cours desquelles le taux d'appels était réparti de la même façon. Ces journées sont les suivantes :

- lundi au jeudi
- vendredi
- samedi et dimanche.

Pour chacune de ces dernières, nous avons estimé le taux d'appels horaire (pour les simulations horizon fini). Le taux d'appels moyen se situe autour de 13 appels/heure, alors que le taux maximal observé est de 21 appels/heure.

- proportions d'appels par secteur

Nous avons localisé les appels sur le réseau agrégé de Montréal (voir [8, p. 21]). Chaque appel a été associé à un noeud de ce réseau. Puisque pour chaque noeud les coordonnées (x,y) sont connues, nous avons pu déterminer le nombre d'appels par secteur* (car chaque secteur est défini par un polygone dont les sommets sont des coordonnées (x,y)).

- priorité

Il existe trois niveaux de priorité, chacun étant défini en terme d'une borne supérieure sur le temps de déplacement. La proportion d'appels pour chaque priorité est donnée par le tableau suivant :

Niveau de priorité	borne supérieure sur le temps de déplacement (en minutes)	proportion
S	5	0.707
2	20	0.231
4	60	0.061

*Note : Dans [14], on parle de 468 secteurs. Nous utilisons maintenant une partition de 484 secteurs.

2) le service

- le type de service (estimation des proportions)

  annulé : 0.215
  non-annulé : 0.785

- loi de probabilité du délai entre l'arrivée sur les lieux et le retour en disponibilité

  Nous n'avons pas pu ajuster la fonction de répartition de ce délai à une loi de probabilité connue. Nous avons donc modélisé les fonctions de répartition empiriques (service annulé et non-annulé) selon la méthode décrite au chapitre 3.

- le C.H. de destination

  Nous n'avons retenu que les 30 C.H. les plus importants. Pour chacun d'eux, nous avons estimé la proportion de services non-annulés. Nous avons également modélisé la loi de probabilité de sélection du C.H. conditionnelle au lieu de l'appel (voir section 4.4).

10.3 La validation

10.3.1 Définitions

Essentiellement, le problème de la validation consiste à se poser la question suivante : est-ce que le modèle de simulation constitue une bonne approximation du système réel? En particulier, est-il possible de s'assurer de la validité du modèle pour une tactique qui est opérationnelle dans la réalité.

Il existe des méthodes pour valider un modèle de simulation d'un système réel en opération. Toutefois, avant de discuter de ces méthodes, nous allons faire une distinction entre la vérification et la validation.

La vérification consiste à s'assurer que le programme de simulation fait bien ce qu'on veut qu'il fasse en utilisant des techniques de "debugging" (construction modulaire, trace, etc.). De plus, pour certaines

tactiques, il est possible d'utiliser un modèle analytique pour vérifier les résultats fournis par le modèle. Nous avons entre autres utilisé le modèle hypercube pour vérifier la tactique garage [1].

Par contre, la validation permet de déterminer si le modèle de simulation (ce qui est différent du programme de simulation) est une bonne approximation du système réel. La validation est un processus qui doit intervenir dans toutes les étapes du développement du modèle.

Comme nous le verrons un peu plus loin, le processus de validation n'a pu être complété entièrement en raison d'absence de données concernant les vitesses des ambulances sur le réseau; ce qui est essentiel pour confronter les résultats de simulation à la réalité. Nous considérons que cette étape devrait éventuellement être complétée afin de s'assurer d'une manière raisonnable de la validité du modèle. Ainsi, nous avons développé une méthodologie pour mener à terme une telle validation. Nous allons présenter les étapes de ce processus en mentionnant celles que nous avons pu accomplir.

10.3.2 Considération sur la validation

Dans la méthode de validation, nous considérons les six (6) points suivants :

1) Le modèle que nous avons développé est une substitution de la réalité; c'est un laboratoire dans lequel on peut essayer différentes tactiques de déploiement. Un des buts est donc de s'assurer que le modèle discrimine les différentes tactiques de la même manière qu'elles le sont dans la réalité.

2) Le modèle est une approximation de la réalité. En validant une tactique particulière, on ne cherche pas à obtenir des performances simulées identiques en tout point à la réalité. On ne veut que se rapprocher suffisamment de cette dernière sans pour autant que les coûts engendrés par le processus de validation soient excessifs.

3) Un modèle valide pour une utilisation particulière ne doit pas être utilisé avec le même degré de confiance pour une utilisation autre que celle pour laquelle il a été validé. Par exemple, le modèle peut avoir été validé pour comparer des tactiques et non pour prévoir les performances absolues.

4) Un modèle est validé selon un ensemble spécifique de critères qui sont habituellement ceux qui interviennent dans la prise de décision (par exemple, le temps de déplacement moyen et $P[T < 5 \text{ min}]$, voir [1]).

5) La validation doit faire partie du processus de développement du modèle.

6) L'utilisation de procédures statistiques pour la validation d'un modèle n'est qu'une partie du processus (quoiqu'importante). En effet, il n'est pas toujours possible d'avoir des données et l'inférence statistique classique n'est pas toujours applicable.

10.3.3 Une approche en trois étapes

Nous considérons une approche en trois étapes de la validation. Cette approche est inspirée de [3] et [23]. Ces trois étapes sont :

- la phase de développement du modèle;
- la vérification de la validité des hypothèses;
- l'évaluation de la représentativité des performances simulées.

1) phase de développement du modèle

a) Le modèle doit être développé en étroite relation avec les experts qui sont familiers avec le fonctionnement du système réel. Nous avons travaillé en consultant chez Urgences santé :

- les gestionnaires chargés de la planification et de l'exploitation des ressources;
- les responsables de la collecte des données pour préciser la nature des données que nous avions en main et compléter les informations manquantes;

- le répartiteur en chef pour définir les tactiques et les protocoles de communication (nous avons également fait de l'écoute parallèle avec les répartiteurs afin de voir de quelle façon ils géraient les opérations);
- des ambulanciers et des superviseurs pour savoir de quelle façon s'enchaînent les différents évènements sur le terrain.

b) On doit collecter le plus d'information possible sur le système. En plus des données décrites en 10.1, nous avons obtenu tous les débuts et fins de quarts (et de pauses) des ambulances dans le mois de juin 1984. Cela est nécessaire pour confronter les performances simulées aux données réelles (étape 3).

c) On doit faire des hypothèses sur le fonctionnement du modèle qui, on l'espère, résisteront à la phase de validation finale. Nous avons entre autres fait les hypothèses suivantes :

- arrivées Poisson des appel;
- les appels n'émanent que des centroïdes de secteurs;
- une seule ambulance est affectée à un appel;
- les ambulances se déplacent entre deux points sur le plus court chemin qui les relie;
- les distances des plus courts chemins sont calculées sur un réseau agrégé;
- la localisation des véhicules en mouvement est faite en utilisant un tableau "approximatif" de successeurs;
- les zones de congestion ont été obtenues arbitrairement;
- les fonctions de vitesse ont été grossièrement approximées.

2) vérification de la validité des hypothèses

- Utilisation de tests statistiques : les hypothèses statistiques ont été vérifiées à l'aide de tels tests (voir [14]).
- S'assurer de la cohérence des informations

On doit éviter le plus possible de développer un modèle pour Urgence santé en utilisant des informations qui proviennent d'une autre ville

ou d'un autre contexte (pompier, par exemple). La plupart de nos informations sur le système proviennent d'Urgences santé sauf pour celles concernant les vitesses.

Note : si on devait obtenir des données pour la calibration, il faudrait s'assurer qu'elles soient prélevées dans des conditions de circulations analogues à celles qui prévalaient au mois de juin 1984.

- Si on voulait diminuer le niveau de détail du modèle (pour des raison d'efficacité), alors on devrait faire une analyse de sensibilité [24]. Par exemple, on pourrait vouloir passer d'une partition de 484 à 200 secteurs. Si le modèle était trop sensible, alors on devrait agréger à un niveau supérieur.

3) évaluation de la représentativité des performances

Cette étape n'est possible que si on dispose d'observations sur les performances du système. On peut alors confronter ces observations aux performances simulées. On prend alors la même séquence d'appels dans les deux cas. On apporte alors les ajustements au modèle jusqu'à ce qu'il y ait une certaine cohérence entre les jeux de données. Nous n'avons pas pu faire cette étape pour les raisons invoquées plus haut. On trouvera dans [3] les détails concernant les techniques pour comparer les performances réelles et simulées.

Dans le présent chapitre, nous avons discuté de la collecte de données et de la validation du modèle pour le cas particulier d'Urgences santé à Montréal. Afin d'implanter le modèle pour d'autres villes ou d'autres contextes (pompier, police), il est évident qu'il faudrait refaire une grande partie du travail effectué pour ces deux aspects.

CHAPITRE 11

CONCLUSION

Dans les systèmes d'urgence, des vies humaines sont en jeu et la santé de la population est directement touchée. Ces systèmes évoluent dans des contextes qui sont complexes et où de multiples phénomènes aléatoires interagissent. L'analyse de ces systèmes n'est pas simple. À cette fin, notre but a consisté à développer un outil de simulation permettant d'évaluer les performances d'un système d'urgence dans le plus large évantail possible de contextes (avec ou sans relocalisation, etc.). Ce but a été atteint avec le modèle développé. Nous avons maintenant un outil parfaitement adapté au cas d'Urgences santé et pouvant être étendu au cas d'un système d'urgence général.

RÉFÉRENCES

[1] TRUDEAU, P., "Planification et opération d'un système d'Urgences santé : la problématique", Centre de recherche sur les transports, Université de Montreal (à paraître).

[2] BRATLEY, P., FOX, B., SCHRAGE, L., "A Guide to Simulation", Springer-Verlag, New York (1983).

[3] LAW, A.M., KELTON, D.W., "Simulation Modeling and Analysis", McGraw-Hill, New York (1982).

[4] FISHMAN, G.S., "Concepts and Methods in Discrete Event Digital Simulation", John Wiley and Sons (1973).

[5] FISHMAN, G.S., "Principles of Discrete Event Simulation", John Wiley and Sons (1978).

[6] HEYMAN, D., SOBEL, M., "Stochastic Process in Operations Research", McGraw-Hill, 1982, (Volume 1).

[7] STANDISH, T.A., "Data Structure Techniques", Addison-Wesley (1980).

[8] ROY, S., ROUSSEAU, J.M., LAPALME, G., FERLAND, J., "Description du système de confection de routes et d'horaires pour le transport adapté Rapport final", Centre de recherche sur les transports, Université de Montréal, publication #413F (1984).

[9] ROY, S., ROUSSEAU, J.M., LAPALME, G., FERLAND, J., "Élaboration des routes et horaires de véhicules pour le transport adapté : l'algorithme", Centre de recherche sur les transports, Université de Montréal, publication #412F (1985).

[10] MALLETTE, C., LAPALME, G., "Guide d'utilisation des Areas Master Files", Centre de recherche sur les transports, Université de Montréal, publication #273 (1982).

[11] Statistique Canada, "Secteurs de recensement, Recensement du Canada de 1981", catalogue 95-91B (1982).

[12] Statistique Canada, Division CANSIM, "CARTLIB 1981, Bibliotèque cartographique en mode numérique des limites des unités statistiques. Secteurs de recensement en 1981".

[13] BRANDEAU, M.L., LARSON, R.C., "Extension and Applying the Hypercube Queing Model to Deploy Ambulances in Boston", Public Systems Evaluation Inc., Cambridge, Mass. (1983).

[14] CHOQUETTE, J., "Sur l'étude du modèle de file d'attente hypercube dans le contexte du déploiement des ambulances à Montréal", Centre de recherche sur les transports, Université de Montréal, publication #450 (1985).

[15] LARSON, R.C., ODONI, A.R., "Urban Operations Research", Prentice-Hall Inc., Englewood Cliffs, N.J. (1981).

[16] WARD, P., "Pascal 6000 Version 3", Centre de calcul, Université de Montréal, bulletin no B60-04, (1980).

[17] CINLAR, E., "Introduction to Stochastic Process", Prentice-Hall, Englewood Cliffs, N.J. (1975).

[18] LAW, A.M., CARSON, J.S., "A Sequential Procedure for Determining the Length of a Steady-State Simulation", Operations Research, 27, 1011-1025 (1979).

[19] GAFARIAN, A.V., ANCKER, C.J., MORISAKU, T., "Evaluation of Commonly Used Rules for Detecting Steady-State in Computer Simulation", Nav. Res. Logist. Q., 25, 511-529, (1978).

[20] SCHRUBEN, L., "Sequential Simulation Run Control Using Standardized Times Series", Cornell University (communication privée) (1986).

[21] TRUDEAU, P., CHOQUETTE, J., FERLAND, J., ROUSSEAU, J-M, "Un modèle de simulation pour un système d'Urgences santé : le guide de l'usager", Centre de recherche sur les transports, Université de Montréal, (à paraître).

[22] MACPHERSON, G., "Les fichiers permanents : une introduction", Centre de calcul, Université de Montréal, bulletin no B01-10B (1984).

[23] NAYLOR, T.H., FINGER, J.M., "Verification of Computer Simulation Models," Management Sciences, 14, 92-101 (1967).

[24] GAFARIAN, A.V., WALSH, J.E., "Methods for Statistical Validation of a Simulation Model for Freeway Traffic Near an On-Ramp", Transportation Research, 4, 379-384 (1970).

[25] TRUDEAU, P., "Comparaison de tactiques d'opérations", Centre de recherche sur les transports, Université de Montréal, (à paraître).